
**Nutzung von
IoT-Technologien und Mobile Computing
zur Verbesserung des
menschlichen Wohlbefindens**

DISSERTATION

zur Erlangung des akademischen Grades
Doktor der Ingenieurwissenschaften (Dr.-Ing.)

der Fakultät für Angewandte Informatik
der Universität Augsburg

vorgelegt von

Andreas Florian SEIDERER

2021

Erstgutachterin:	Prof. Dr. Elisabeth André
Zweitgutachter:	Prof. Dr. Björn Schuller
Tag der mündlichen Prüfung:	23.07.2021

Abstract

Human well-being is positively and negatively influenced by many factors. With the ongoing expansion of ubiquitous systems which can be implemented with Internet of Things (IoT) and Mobile Computing technologies, networked devices with a variety of sensors and actuators are increasingly integrated into the environment or worn by people. These technologies therefore have the potential to acquire and process relevant factors for well-being automatically or by involving the user and to output the results in various forms that should be as comprehensible as possible for the respective user.

The aim of this work is to identify possibilities of using IoT- and Mobile Computing technologies to improve human well-being to develop and investigate methods as well as assistance, information and recommendation systems suitable for everyday use. To achieve this goal, a pragmatic well-being model was developed that included the five categories emotional, mental, social, environmental and physical. For each category, relevant well-being factors based on interdisciplinary literature were identified and investigated how to capture them by IT systems. The acquisition of a selection of these factors is addressed more comprehensively in studies using specific types of unobtrusive sensors, such as gas sensors and pressure-sensitive mats, and presenting solutions for real-time signal processing using various machine learning techniques, such as deep neural networks and transfer learning. To enhance the user's emotional, mental and social wellbeing, the PiHeart system was implemented, a biofeedback system with a tangible interface that enables users to experience heart beats as natural stimuli from themselves and another person in social proximity. Furthermore, to improve environmental well-being, two information systems were developed that use ambient light for unobtrusive system output. One of the systems provides an overview of the opening status of safety-critical windows and doors in the home, while the other is used to assess indoor air quality. To improve the physical well-being of users, the Drink-O-Mender system was developed, in which a social robot provides recommendations regarding healthy drinks. Furthermore, a speech assistant in the German language is presented that is able to answer questions regarding nutritional information. Finally, the CARE system, a context-aware lifestyle recommender system for older adults is introduced. It includes all categories of the well-being model. CARE chooses the currently most suitable recommendations based on data acquired by sensors embedded in the user's home environment. For validation, several hardware and software prototypes were developed and tested with different user groups, not only under laboratory conditions, but also in everyday life. To ensure enduring user acceptance, the systems were developed with a focus on high usability, taking into account, among other things, system usage data as well as feedback from semistructured interviews and workshops with potential users of the target groups. Furthermore, privacy protection was ensured for the systems by investigating and applying methods for local data processing on user devices, e. g., the use of transfer learning with deep neural networks on mobile devices. In this work particular emphasis was given on the question how to smoothly integrate technologies from the IoT and Mobile Computing fields into users' daily lives to improve their well-being. To this end, various interaction techniques were implemented that take place in the users' background of attention and thus don't disturb them during their daily activities.

Zusammenfassung

Das menschliche Wohlbefinden wird von vielen verschiedenen Faktoren positiv und negativ beeinflusst. Durch die fortschreitende Verbreitung von ubiquitären Systemen, die mit Technologien des Internet of Things (IoT) und Mobile Computings umgesetzt werden können, werden zunehmend vernetzte Geräte mit einer Vielzahl an Sensoren und Aktuatoren in die Umgebung integriert oder von Personen getragen. Diese Technologien besitzen daher das Potenzial, automatisch oder durch Einbeziehung des Nutzers, relevante Faktoren für das Wohlbefinden zu erfassen, zu verarbeiten sowie die Ergebnisse in verschiedenartiger Form, die für den jeweiligen Nutzer möglichst gut verständlich sein soll, auszugeben.

Das Ziel dieser Arbeit ist, Möglichkeiten der Nutzung von Technologien aus dem IoT- und Mobile-Computing-Bereich zur Verbesserung des menschlichen Wohlbefindens zu identifizieren, geeignete Verfahren sowie alltagstaugliche Assistenz-, Informations- und Empfehlungssysteme zu entwickeln und zu untersuchen. Um dieses Ziel zu erreichen, wurde ein pragmatisches Wohlbefindensmodell entwickelt, das die fünf Kategorien emotional, mental, sozial, umgebungsbezogen und physisch beinhaltet. Für jede Kategorie werden zunächst maßgebliche Wohlbefindensfaktoren identifiziert, die auf aktueller wissenschaftlicher Literatur aus verschiedenen Disziplinen basieren, und von IT-Systemen erfasst werden können. Die Erfassung einer Auswahl dieser Faktoren wird in Untersuchungen umfassender behandelt, wobei spezielle Arten unaufdringlicher Sensorik, wie Gassensoren und drucksensitive Matten eingesetzt und Lösungen zur Echtzeitsignalverarbeitung unter Verwendung verschiedener Verfahren des maschinellen Lernens, wie z. B. tiefen neuronalen Netzen und Transferlernen, vorgestellt werden. Um das emotionale, mentale und soziale Wohlbefinden von Nutzern zu verbessern, wurde das PiHeart-System entwickelt. Es handelt sich hierbei um ein Biofeedback-System, das den eigenen Herzschlag bzw. den einer anderen Person mittels eines Tangible-Displays spürbar macht. Des Weiteren wurden zur Verbesserung des umgebungsbezogenen Wohlbefindens zwei Informationssysteme entwickelt, die ambientes Licht zur unaufdringlichen Systemausgabe nutzen. Eines der Systeme ermöglicht einen Überblick über den Öffnungszustand von sicherheitskritischen Fenstern und Türen in der Wohnung, das andere dient der Bewertung der Raumluftqualität. Um das physische Wohlbefinden von Anwendern zu verbessern, wurde das Drink-O-Mender-System entwickelt, bei dem ein sozialer Roboter Empfehlungen bezüglich gesunder Getränke gibt. Des Weiteren wird ein deutschsprachiger Sprachassistent präsentiert, der Fragen hinsichtlich Nährwertinformationen von Lebensmitteln beantworten kann. Abschließend wird das CARE-System, ein kontextsensitives Lifestyle-Empfehlungssystem für Senioren, vorgestellt. Es beinhaltet alle Kategorien des Wohlbefindensmodells. Die für einen Nutzer aktuell geeignetsten Empfehlungen werden, basierend auf Daten von in der Wohnumgebung integrierten Sensorik, ausgewählt. Zur Validierung wurden mehrere Hardware- und Software-Prototypen entwickelt und mit verschiedenen Nutzergruppen, nicht nur unter Laborbedingungen, sondern auch im Alltag getestet. Um eine nachhaltige Nutzerakzeptanz zu gewährleisten, wurde bei der Entwicklung der Systeme auf eine hohe Usability geachtet, wobei u. a. Daten der Systemnutzung sowie Feedback aus semistrukturierten Interviews und Workshops mit potenziellen Anwendern der Zielgruppen beim Entwicklungsprozess berücksichtigt wurden. Des Weiteren wurde der Schutz der Privatsphäre bei den Systemen gewährleistet, indem Verfahren zur lokalen Datenverarbeitung auf Nutzergeräten, z. B. der Einsatz von Transferlernen mit tiefen neuronalen Netzen auf Mobilgeräten, untersucht und angewandt wurden. Besonderes Augenmerk wurde in dieser Arbeit auf die Frage gelegt, wie sich Technologien aus dem IoT- und Mobile-Computing-Bereich zur Verbesserung des Wohlbefindens von Nutzern reibungslos in ihren Alltag integrieren lassen. Hierzu wurden verschiedene Interaktionstechniken implementiert, die im Aufmerksamkeitshintergrund der Nutzer stattfinden und sie somit nicht bei ihren täglichen Aktivitäten stören.

Danksagung

Mein besonderer Dank gilt meiner Doktormutter Prof. Dr. Elisabeth André, für die sehr gute Zusammenarbeit und Unterstützung bei der Erstellung dieser Arbeit. Zudem bedanke ich mich bei Prof. Dr. Björn Schuller für dessen Bereitschaft diese Arbeit zu begutachten und bei Prof. Dr. Jörg Hähner, der sich als Prüfer bereit erklärte.

Des Weiteren möchte ich mich bei meinen aktuellen und ehemaligen Kollegen am Lehrstuhl, insbesondere bei Hannes Ritschel, Simon Flutura, Dr. Chi Tai Dang, Dr. Ilhan Aslan, Dr. Stephan Hammer und Dr. Johannes Wagner für die Zusammenarbeit bedanken. Für die sehr gute Kooperation im Rahmen des CARE-Projekts spreche ich Prof. Dr. Thomas Rist und Madita Herpich meinen Dank aus.

Anmerkungen

- Für den physikalischen Begriff „elektrische Arbeit“ werden in einer Vielzahl an Fachliteratur und in der Umgangssprache die Begriffe „Stromverbrauch“ oder „Energieverbrauch“ synonym verwendet. Diese gebräuchlicheren Begriffe werden daher auch in dieser Arbeit genutzt.
- In dieser Arbeit wird der APA-Zitierstil verwendet. Innerhalb der Arbeit werden Abschnitte mit „siehe“ bzw. „siehe auch“ referenziert. Weiterführende Literatur wird mit „siehe hierzu“ zitiert.
- Das Aufrufdatum von weiterführenden Internetlinks wird in Klammern angegeben.
- Zur besseren Lesbarkeit wird in dieser Arbeit das generische Maskulinum verwendet. Falls im Kontext kein Geschlecht angegeben wurde, sind weibliche und anderweitige Geschlechteridentitäten dabei ausdrücklich mitgemeint.

Inhaltsverzeichnis

Zusammenfassung	v
I Einführung	1
1 Einführung	3
1.1 Ubiquitous Computing, IoT und Mobile Computing	5
1.2 Privatsphäre	5
1.3 Herausforderungen und Forschungsziele	6
1.4 Struktur der Arbeit	8
II Konzeptuelle und technische Grundlagen	11
2 Wohlbefindensmodell	13
3 IoT und Mobile Computing: Plattformen und Sensorik	17
3.1 Mikrocontroller und Einplatinencomputer	17
3.2 Mobile Geräte	19
3.2.1 Systeme	19
3.2.2 Mobile Datenverarbeitung	20
3.3 Smarthome	23
3.4 Verwendete Sensorik	24
3.5 Überblick Gassensorik	30
3.5.1 Messprinzipien	30
3.5.2 Sensorpositionierung	31
3.5.3 Herausforderungen bei der Auswertung	32
3.5.4 Verwendete Sensoren	33
3.5.5 Prototypen	36
4 Echtzeitsignalverarbeitung und maschinelles Lernen	39
4.1 Kontinuierliche Daten	39
4.1.1 Das Social Signal Interpretation (SSI) Framework	40
4.1.2 MobileSSI	41
4.2 Diskrete / eventbasierte Daten	41
4.2.1 Hausautomations-Software	42
4.2.2 Eventerpretor	43
4.3 Maschinelles Lernen	45
4.3.1 Eingesetzte Verfahren	45
4.3.2 Klassifizierung	46

5	Transferlernen	49
5.1	Definitionen für Domäne, Aufgabe und Transferlernen	50
5.2	Probleme beim Transferlernen	50
5.3	Kategorien von Transferlernen	51
5.3.1	Induktives Transferlernen	51
5.3.2	Unüberwachtes Transferlernen	52
5.3.3	Transduktives Transferlernen	52
5.4	Ansätze beim Transferlernen	53
5.5	Transferlernen beim Deep Learning	53
5.5.1	Wissenstransfer mittels Netzwerk	54
5.5.2	Untersuchung der Leistungsfähigkeit von Transferlernen	54
5.5.3	Nutzung von Audiospektrogrammen	61
5.5.4	Anwendungsbeispiele für Transferlernen	62
III	Verfahren und Systeme zur Verbesserung des menschlichen Wohlbefindens	63
6	Emotionales, mentales und soziales Wohlbefinden	65
6.1	Erfassung von Stress, Lachen und sozialer Verbundenheit	66
6.1.1	Stress und dessen Auswirkungen auf das Wohlbefinden	66
6.1.2	Lachen als positiver Emotionsausdruck und dessen Auswirkungen auf das Wohlbefinden	67
6.1.3	Soziale Verbundenheit und deren Auswirkungen auf das Wohlbefinden . . .	68
6.1.4	Eigene Arbeiten	68
6.2	Biofeedback mittels Tangible-Displays	69
6.3	Eigene Untersuchungen	70
6.3.1	Stresserkennung bei der Arbeit	70
6.3.2	Erfassung von Lachen mit Smartphones	79
6.3.3	Biofeedback des Herzschlags durch Tangible-Displays zur Verstärkung des eigenen und geteilten affektiven Erlebnisses	82
6.4	Zusammenfassung des Kapitels	93
7	Umgebungsbezogenes Wohlbefinden	95
7.1	Erfassung von Umwelteinflüssen	97
7.1.1	Umgebungsparameter und deren Auswirkungen auf das Wohlbefinden . . .	97
7.1.2	Eigene Arbeiten	102
7.2	Ambiente Systeme	103
7.3	Eigene Untersuchungen	105
7.3.1	Klassifizierung von Klanglandschaften mit generierten DNNs	105
7.3.2	Auswirkungen der städtischen Umgebung auf den Menschen	113
7.3.3	Ambiente Anzeige des Status von Fenstern / Türen zur Verbesserung des Sicherheitsempfindens in der Wohnumgebung	120
7.3.4	Verbesserung der Luftqualität in einem Besprechungsraum	128
7.4	Zusammenfassung des Kapitels	139
8	Physisches Wohlbefinden: Ernährung	141
8.1	Nahrungserfassung	143
8.1.1	Erfassung mittels Mobile Computing und IoT	143
8.1.2	Eigene Arbeiten	144
8.2	Abfrage von Informationen zu Lebensmitteln	146
8.2.1	Interaktionsgeräte	146

8.2.2	Eigene Arbeiten	147
8.3	Eigene Untersuchungen	148
8.3.1	Erkennung von Lebensmitteln beim Sprechen während des Verzehrens mittels Audiodaten	148
8.3.2	Erkennung von Lebensmitteln mittels Gassensorik	152
8.3.3	Erkennung von Lebensmitteln mittels Gassensorik: Transferlernen zur Kompensation des einzigartigen Sensorverhaltens	157
8.3.4	Vereinfachung der mobilen Ernährungserfassung mittels Interaktion mit Smartscale, Smartwatch und Smartphone	160
8.3.5	Erkennung von Trinkvorgängen mittels einer Smartwatch und interaktivem / kooperativem Lernen	174
8.3.6	Entwicklung eines adaptiven Roboterassistenten für Informationen über Getränke	181
8.3.7	Entwicklung eines deutschsprachigen Sprachassistenten für Senioren zur Abfrage von Informationen über Lebensmittel	189
8.4	Zusammenfassung des Kapitels	198
9	Projekt CARE: Verbesserung des Wohlbefindens von Senioren	201
9.1	Konzept CARE-System	202
9.1.1	Anforderungen der Zielgruppe Senioren	202
9.1.2	Verwandte Literatur	205
9.1.3	Ergebnisse	207
9.2	Empfehlungen zur Verbesserung des Wohlbefindens	207
9.2.1	Modell des Wohlbefindens für CARE	207
9.2.2	Umsetzung	208
9.3	Sensorik für das CARE-System	210
9.3.1	Stationäre Sensorik	210
9.3.2	Mobile Sensorik	212
9.4	Entwicklung CARE-Prototyp 1	213
9.4.1	Umsetzung	214
9.4.2	Evaluation	220
9.5	Entwicklung CARE-Prototyp 1.5 und 2	226
9.5.1	Weiterentwicklungen	226
9.5.2	Umsetzung	226
9.5.3	Evaluation	234
9.5.4	Zusammenfassung	236
9.6	Entwicklung CARE-Prototyp 2 mit Gamifizierungen	236
9.6.1	Gamifizierung	236
9.6.2	Workshops	237
9.6.3	Umsetzung im CARE-System	239
9.6.4	Evaluation	240
9.6.5	Zusammenfassung	241
9.7	Roboterassistent für Senioren	241
9.7.1	Einleitung	241
9.7.2	Systemaufbau	244
9.7.3	Evaluation	245
9.7.4	Zusammenfassung	247
9.8	Zusammenfassung des Kapitels	247

IV Zusammenfassung	249
10 Zusammenfassung und Ausblick	251
10.1 Wissenschaftliche Beiträge	252
10.1.1 Entwicklung von Verfahren zur Erfassung von Wohlbefindensfaktoren	252
10.1.2 Entwicklung geeigneter Computersysteme zur Verbesserung des Wohlbefindens	256
10.1.3 Gewährleistung der Privatsphäre	259
10.2 Fortführende Arbeiten	260
A Publikationen und Reviewaktivitäten	261
A.1 Wissenschaftliche Veröffentlichungen	261
A.2 Durchgeführte Reviews	267
B Lehre	268
B.1 Betreute Lehrveranstaltungen	268
B.2 Betreute Abschlussarbeiten	268
C (mobile)SSI Plugins	269
Literatur	270

Teil I

Einführung

1	Einführung	3
1.1	Ubiquitous Computing, IoT und Mobile Computing	5
1.2	Privatsphäre	5
1.3	Herausforderungen und Forschungsziele	6
1.4	Struktur der Arbeit	8

1 Einführung

Seit einigen Jahren nimmt die Verbreitung von Smarthomes und mobilen Geräten stetig zu. Dies liegt technologisch daran, dass immer leistungsfähigere und zugleich kostengünstige vernetzte Hardware mit entsprechender Software verfügbar ist, die zu Internet-of-Things- und Mobile-Computing-Technologien zählen. Die Kombination dieser Technologien erlaubt es, ubiquitäre Systeme zu entwickeln. Smarthome-Systeme, die aus verschiedenen vernetzten Sensor- und Aktor-Geräten bestehen, werden beispielsweise laut der „Smart Home Studie 2020“ von Bitkom Research¹ häufig zur Automatisierung der Beleuchtung, der Erhöhung der Sicherheit z. B. durch Alarmierung bei Einbrüchen oder für eine energieeffiziente Heizungssteuerung eingesetzt. Auf diese Weise können sie u. a. den Komfort verbessern sowie für ein erhöhtes Gefühl von Sicherheit und angenehmes Raumklima sorgen, wodurch das Wohlbefinden der Bewohner gesteigert werden kann. Seit kürzerer Zeit sind zudem kommerzielle Sprachassistenten verfügbar, die gemäß der Bitkom-Studie bereits von ca. 50 % der Smarthome-Nutzer verwendet werden. Alexa, Google Assistant oder Siri stellen das aktuellste Interface für Smarthomes dar und können wesentlich erweiterte Funktionen anbieten, indem sie zusätzlich Internetdienste zur Verarbeitung und Speicherung von Daten einbeziehen und eine intuitive Sprachbedienung mittels Smartspeakers ermöglichen, die teilweise auch über Displays verfügen. Sprachassistenten können beispielsweise per Sprach-eingabe Internetsuchen durchführen, Musik abspielen, Termine in den persönlichen Kalender eintragen, Kochrezepte vorlesen oder das Licht im Wohnzimmer steuern. Sprachassistenten sind zudem auf mobilen Geräten verfügbar und können daher von nahezu überall verwendet werden. Obwohl durch den Zusammenschluss des Smarthomes und mobiler Geräte potenziell eine große Anzahl an Kontextinformationen zur Verbesserung des Wohlbefindens der Nutzer verfügbar wären, existieren hierfür derzeit nur wenige Anwendungen. Auf diese Thematik spezialisierte Smartphone-Apps sind jedoch in einer großen Anzahl in den Appstores aufzufinden. Sie decken u. a. gesunde Ernährung oder sportliche Fitness ab. Allerdings sind diese Apps meist auf Nutzereingaben und die Sensordaten des Smartphones oder eines Fitnessarmbands beschränkt. Insbesondere für Empfehlungs- und Assistenzsysteme wäre es jedoch von großem Vorteil, auf möglichst viele Kontextinformationen aus dem Smarthome und von mobilen Geräten zugreifen zu können.

Das Ziel dieser Arbeit ist, Möglichkeiten zur Nutzung von Technologien aus dem IoT- und Mobile-Computing-Bereich zur Verbesserung des menschlichen Wohlbefindens zu identifizieren, entsprechend angepasste Verfahren sowie Assistenz-, Informations- und Empfehlungssysteme zu entwickeln und zu untersuchen. Hierzu ist es erforderlich, positive und negative Faktoren quantitativ zu bestimmen, die auf das menschliche Wohlbefinden Einfluss haben. In der Forschung wurden hierfür, wie z. B. beim BeWell-System von Lane et al. (2014) vereinfachte Modelle eingesetzt, um beispielsweise die physische Aktivität als Teil des Wohlbefindens quantitativ zu erfassen. In dieser Arbeit wird ein umfassenderes Wohlbefindensmodell entwickelt, das auf bereits etablierten Modellen zur Ermittlung der Lebensqualität basiert. Das Modell wird in einem kontextsensitiven Empfehlungssystem speziell für Senioren eingesetzt.

Insbesondere Senioren könnten von Assistenz-, Informations- und Empfehlungssystemen profitieren, die sie bei der Verbesserung ihres Wohlbefindens unterstützen. In Kombination mit anderen Technologien aus dem Ambient-Assisted-Living-Bereich kann es ihnen auch dabei helfen, länger in ihrer eigenen Wohnung leben zu können. Allerdings sind aktuell die wenigsten kommerziellen Smarthome-Systeme auf diese Zielgruppe abgestimmt. Senioren sind zudem in der Regel

¹[https://www.bitkom.org/Bitkom/Publikationen/Smart-Home-Studie-2020\(03.05.2021\)](https://www.bitkom.org/Bitkom/Publikationen/Smart-Home-Studie-2020(03.05.2021))

weniger technikaffin als jüngere Altersgruppen. Dies wurde auch in der Bitkom-Studie 2020 festgestellt, bei der nur 13 % der befragten über 65-Jährigen Smarthome-Anwendungen nutzten. Bei dieser Altersgruppe war zudem eine hohe Usability das wichtigste Kriterium beim Kauf von Smarthome-Anwendungen. Sie nannten eine zu komplizierte Bedienung als Hauptgrund, bisher keine Smarthome-Anwendungen zu nutzen. In dieser Arbeit werden daher in einigen Untersuchungen die Anforderungen von Senioren an IT-Systeme ermittelt und in einem nutzerzentrierten Entwicklungsprozess berücksichtigt.

Generell stellt ein aufdringliches Systemverhalten bei der Interaktion mit Systemen im Alltag ein großes Problem dar, da es deren Nutzerakzeptanz reduziert. Die Systeme sollten sich daher möglichst nahtlos in die Alltagsroutinen der Nutzer einfügen. Eine Lösungsmöglichkeit besteht darin, ambiente Systeme zu entwickeln bei denen die Grundidee, ubiquitärer Systeme im Hintergrund zu agieren, praktisch angewandt wird. Ambiente Systeme können in die Umgebung integrierte Sensorik, wie beispielsweise Gassensorik, einsetzen und ambiente Anzeigen, wie z. B. ambientes farbiges Licht, zur Systemausgabe nutzen. Ein Großteil der Smarthome-Hardware ist als Basis für ambiente Systeme bereits grundsätzlich geeignet. Darauf basierende ambiente Informationssysteme werden in dieser Arbeit weitergehend untersucht. Ein weiterer Teil des Lösungsansatzes zur Vermeidung aufdringlicher Systeme stellt eine erleichterte Erfassung von Wohlbefindensfaktoren dar. Im Idealfall können diese automatisch im Hintergrund ermittelt werden. Dies ist beispielsweise unter Verwendung von Sensordaten stationärer und mobiler Geräte mit entsprechenden Verfahren der Echtzeitsignalverarbeitung und des maschinellen Lernens möglich. Einige Informationen können jedoch nicht ohne Nutzerinteraktion ermittelt werden, da entweder keine geeignete Sensorik vorhanden oder die Präzision automatischer Erfassungstechniken (aktuell) zu gering oder sehr fehleranfällig ist. In diesem Fall sollte eine hohe Usability gewährleistet werden. In dieser Arbeit werden für verschiedene Wohlbefindensfaktoren automatische Erfassungsmethoden untersucht, die von Assistenz-, Informations- und Empfehlungssystemen eingesetzt werden könnten. Hierbei werden spezielle Sensoren, wie drucksensitive Matten und Gassensoren untersucht, die bislang primär in der Forschung eingesetzt werden und vielversprechende Ergebnisse zeigten. Insbesondere bei der Verwendung spezieller Sensorik sind allerdings in der Regel nur wenige Datenkorpora verfügbar und es können nur vergleichsweise geringe Datenmengen für Klassifizierungen aufgezeichnet werden. Besonders problematisch ist dies, wenn leistungsfähige tiefe neuronale Netze zur Anwendung kommen, die große Datenmengen zum Trainieren benötigen. In der Forschung hat sich bereits für viele Anwendungszwecke gezeigt, dass hierfür Transferlernen (siehe hierzu den Literaturüberblick von K. Weiss et al. (2016)) geeignet ist und diese Methodik wird folglich auch in dieser Arbeit untersucht. Bei tiefen neuronalen Netzen besteht zudem das Problem der Entwicklung optimaler Netzarchitekturen. Genetische Algorithmen können hierbei eine Lösung darstellen, wenn diese tiefe neuronale Netze mit akzeptabler Leistung erzeugen können und eine annehmbare Laufzeit besitzen. Eine Reduktion des Aufwandes bei der Erfassung von Wohlbefindensfaktoren, an denen der Nutzer beteiligt sein muss, kann mit mehreren Methoden erreicht werden. Die nachfolgend vorgestellten Methoden werden daher in dieser Arbeit untersucht. Insbesondere bei der Personalisierung von Klassifikationsmodellen kann interaktives maschinelles Lernen einen geeigneten Lösungsansatz darstellen, da durch eine abnehmende Anzahl an Nutzerabfragen die Klassifizierung kontinuierlich verbessert werden kann. Da Nutzer in vielen Fällen bereits mehrere mobile Geräte einsetzen, wie Smartphone und Smartwatch, könnten diese von Nutzern situationsabhängig im Rahmen einer Multi-Geräte-Interaktion zur Erfassung eingesetzt werden.

Generell stellt insbesondere bei gesundheitsbezogenen Systemen die Wahrung der Privatsphäre ein aktuelles Problem dar. In der Bitkom-Studie von 2020 nannten 33 % bzw. 24 % der befragten Personen, dass sie bisher keine Smarthome-Anwendungen nutzen, da sie Angst vor Missbrauch ihrer persönlichen Daten bzw. Angst um ihre Privatsphäre besitzen. Insbesondere durch die Datenverarbeitung durch Cloud-Dienste ist diese Angst nicht unbegründet, da in vielen Fällen Sensordaten aus dem Smarthome oder von Fitnessarmbändern permanent im Hintergrund an

die Anbieter übermittelt werden, teilweise ohne, dass sich Nutzer darüber bewusst sind. Es ist allerdings kaum möglich zu überprüfen, wie ein Anbieter die Daten nutzt. Es stellt sich daher die Frage, inwiefern Daten auch mit komplexeren Auswertungsverfahren auf den Geräten der Nutzer verarbeitet werden können. Ein Lösungsansatz wäre lokale Echtzeitsignalverarbeitung auf mobilen und Embedded Systemen einzusetzen. Beim maschinellen Lernen stellt wiederum Transferlernen zum Training tiefer neuronaler Netze einen vielversprechenden Ansatz dar. Besonders kritisch sind auch Audiodaten, die von Sprachassistenten an Cloud-Dienste übertragen werden, die auch lokal verarbeitet werden könnten.

1.1 Ubiquitous Computing, IoT und Mobile Computing

Weiser (1991) stellte in seinem Artikel die Prognose auf, dass im 21. Jahrhundert Computer allgegenwärtig (ubiquitär; engl. ubiquitous) sein werden und im Hintergrund agieren. Er bezeichnete dies als **Ubiquitous Computing**. Die Geräte wären verschlüsselt miteinander vernetzt und könnten durch ihre Hintergrundaktivität verhindern, dass eine Informationsüberflutung beim Menschen stattfindet. Teile dieser Vision sind heutzutage Realität und werden mittels des Internets der Dinge und Mobile Computing umgesetzt. Sie automatisieren und erleichtern damit viele Aufgaben in unserem Umfeld. Klassische Arbeitsplatzcomputer sind allerdings nicht verschwunden und auch heute noch wird mit vielen Systemen im Vordergrund interagiert – sie befinden sich somit im Zentrum der Aufmerksamkeit. In einigen Fällen findet ein fließender Übergang der Interaktion zwischen Vordergrund und Hintergrund statt. Eine Einordnung entsprechender Systeme ist mit dem *Implicit Interaction Framework* von Ju und Leifer (2008) möglich.

Mit Technologien des **Internets der Dinge** werden heutzutage viele Geräte miteinander vernetzt, die mit dem Internet verbunden sind. Der Begriff „Internet of Things“ wurde zuerst 1999 von Kevin Ashton in einer Präsentation genannt (Rayes & Salam, 2019). Es existiert jedoch keine einheitliche Definition. Gemäß Rayes und Salam (2019) ist die Hauptidee von IoT, physikalisch alle Dinge miteinander über das Internet zu verbinden, um sie zu überwachen oder zu steuern. Dinge können hierbei Geräte, Gebäude oder Autos, aber auch Menschen, Tiere oder Pflanzen sein. Die Autoren definieren IoT sinngemäß als das Netzwerk von Dingen, das Menschen und Dinge über das Internet miteinander verbindet. Das Netzwerk besteht hierbei aus Dingen, die eine eindeutige Geräteidentifikation, eingebettete Intelligenz sowie die Möglichkeit besitzen, die Umgebung zu erfassen und zu agieren. Es ist hierbei hervorzuheben, dass diese Definition von IoT auch eine Kommunikation zwischen Mensch und Maschine einschließt. Wichtige Komponenten des Internets der Dinge sind unter anderem Sensoren, Aktuatoren, Identifikatoren für Geräte und Datenquellen, Software zur Datenanalyse und die Fähigkeit, eine Internetverbindung herzustellen. Ein Anwendungsfall von IoT stellt das Smarthome dar. Smarthome-Komponenten könnten daher für in den Alltag integrierte, unaufdringliche Systeme zur Verbesserung des Wohlbefindens sehr gut geeignet sein.

Damit auch unterwegs für das Wohlbefinden relevante Daten erfasst werden können, wird mobile Hard- und Software aus dem **Mobile-Computing**-Bereich benötigt und ist daher auch für diese Arbeit von Relevanz.

1.2 Privatsphäre

Sowohl in der deutschen Verfassung als auch in der Charta der Grundrechte der Europäischen Union wurde der Schutz der Privatsphäre als Menschenrecht verankert. Dies betrifft insbesondere den häuslichen Lebensbereich, aber auch die allgemeine persönliche Integrität als Teil einer freien Lebensführung. Hieraus ergeben sich besondere Herausforderungen für den Einsatz von IT-Systemen.

IoT und Mobile Computing Insbesondere in den letzten Jahren wurde das Thema „Schutz der Privatsphäre“ besonders präsent, da eine immer größere Verbreitung an IoT- und mobilen Geräten mit umfangreicher Sensorik erfolgt, die dauerhaft mit dem Internet und Anbietern von Online-Diensten verbunden sind. Hierbei musste festgestellt werden, dass sehr häufig von Anbietern die Datenschutzbestimmungen nicht eingehalten wurden und dieses durch Medien veröffentlichte Fehlverhalten das Vertrauen in eine sichere Handhabung und Verwahrung persönlicher Daten stark beeinträchtigte. Hierzu zählen beispielsweise die unerwünschte Aufzeichnung und Auswertung von Audiodaten bei Sprachassistenten (z. B. online Artikel der Süddeutschen Zeitung vom 23.10.2019²) oder die Erfassung von Bewegungsprofilen mittels Smartphones (z. B. Artikel auf heise.de vom 07.05.2011³). Viele Anbieter betonen, dass sie diese Daten zur Verbesserung ihrer Dienste verwenden und dies vorteilhaft für den Nutzer sei. Allerdings ist es in vielen Fällen unklar, was mit den Daten im Detail erfolgt und inwieweit Anwender diese Problematik verstehen, wenn sie den Allgemeinen Geschäftsbedingungen einer Firma zustimmen. Dies wird zusätzlich verschärft, wenn es sich um gesundheitsbezogene Daten handelt. Folglich ist im Rahmen dieser Arbeit hoher Wert auf den Schutz der Privatsphäre zu legen und dies bei der Entwicklung von IT-Systemen von Beginn an zu beachten.

Privatsphäre bei der Datenverarbeitung Vor allem bei gesundheitsbezogenen IoT-Anwendungen, bei denen besonders schützenswerte Daten verarbeitet werden, stellt der Schutz der Privatsphäre eine große Herausforderung dar. Dies ist daher auch bei dieser Arbeit von Relevanz. Im Buch „Handbook of Multimodal-Multisensor Interfaces Volume 3“ von Oviatt et al. (2019) werden in Kapitel 16.3 verschiedene Arten von Datenschutzproblemen mit strukturierten Daten und mögliche Lösungsansätze zur Anonymisierung aufgezeigt. In Kapitel 16.4 werden verschiedene Techniken zur Deanonymisierung und Bestimmung des Aufenthaltsortes erwähnt, die mit verschiedenen Sensordaten durchgeführt werden können. Hierzu zählen die Bestimmung des Ortes einer Person durch Bilder oder Ton bzw. die Deanonymisierung durch Nutzung von Sensoren mit einzigartigem Verhalten, z. B. dem Bildrauschen einer Kamera. Diese Beispiele zeigen, dass auch aus offensichtlich unkritischen Daten Informationen gewonnen werden, welche die Privatsphäre verletzen können. Auch mittels der Daten berechnete Merkmale ermöglichen in vielen Fällen eine (verlustbehaftete) Rekonstruktion der Originaldaten (z. B. (inverse) Fouriertransformation). Bei Klassifizierungsmodellen besteht ebenfalls die Möglichkeit, Rückschlüsse auf die Trainings-Daten zu ziehen. Beispielsweise können bei tiefen neuronalen Netzen mittels gezielter Inferenzattacken Daten abgegriffen werden, die beim Trainieren genutzt wurden (Nasr et al., 2019).

1.3 Herausforderungen und Forschungsziele

In dieser Arbeit werden im Hinblick auf die Entwicklung bzw. Nutzung von IT-Systemen zur Verbesserung des menschlichen Wohlbefindens spezielle Verfahren zur Identifikation und Erfassung von Wohlbefindensfaktoren behandelt. Die damit verbundenen wissenschaftlichen und technischen Herausforderungen sowie Forschungsziele werden in diesem Abschnitt vorgestellt.

Identifikation von relevanten Wohlbefindensfaktoren

Für die Entwicklung von Verfahren und Systemen zur Verbesserung des menschlichen Wohlbefindens ist primär ein quantifizierbares Wohlbefindensmodell erforderlich, das Faktoren enthält, die in Kategorien aufgeteilt sein können. Die Faktoren der jeweiligen Kategorien sollten durch implizite oder explizite Nutzereingaben und durch Sensorik erfassbar sein. Grundsätzlich geeignete Wohlbefindensmodelle wurden in der Literatur bereits für verschiedene Anwendungen, wie z. B.

²<https://www.sueddeutsche.de/digital/amazon-echo-google-home-datenschutz-1.4650784> (28.06.2020)

³<http://heise.de/-1236861> (28.06.2020)

für Fragebögen, entwickelt und evaluiert. Hieraus abgeleitete Wohlbefindensfaktoren können aus unterschiedlichen Disziplinen, wie Medizin, Psychologie und Geografie, stammen.

Ein Ziel dieser Arbeit besteht darin, ein Wohlbefindensmodell, basierend auf bereits etablierten Modellen, zu entwickeln und hierfür adäquate Wohlbefindensfaktoren zu identifizieren, die mittels IoT- und Mobile-Computing-Technologien u. a. in der Wohnumgebung und mobil erfasst werden können. Die identifizierten Faktoren sollten mit geeigneten Verfahren bestimmt und in Assistenz-, Informations- und Empfehlungssystemen zur Verbesserung des Wohlbefindens praktisch nutzbar sein. Da in dieser Arbeit der Schutz der Privatsphäre hohe Priorität haben sollte, musste dies bereits von Beginn an bei der Konzepterstellung berücksichtigt werden. So sollte auch bei der Erfassung der Wohlbefindensfaktoren darauf geachtet werden, auf für die Privatsphäre potenziell kritische Sensorik, wie Videokameras, zu verzichten. Die Identifikation von Wohlbefindensfaktoren anhand der zuvor genannten Anforderungen stellt eine Besonderheit im Vergleich zu bereits veröffentlichten Arbeiten dar.

Entwicklung von Verfahren zur Erfassung von Wohlbefindensfaktoren

Um die identifizierten Wohlbefindensfaktoren erfassen zu können, ist es erforderlich, angepasste Verfahren zu entwickeln und untersuchen. Hierfür ist neben der Auswahl geeigneter Sensorik entsprechende Datenverarbeitung notwendig, um benötigte Informationen zu extrahieren. In verschiedenen Forschungsbereichen der Informatik können bereits Arbeiten gefunden werden, die sich potenziell für eine Erfassung bestimmter Wohlbefindensfaktoren anbieten. Zu den Forschungsgebieten zählen beispielsweise Affective Computing, Ambient Assisted Living oder Human Activity Recognition. In diesen Gebieten wurden bereits Erfassungstechniken entwickelt, die teilweise Echtzeitsignalverarbeitung und maschinelles Lernen einsetzen, um u. a. automatisch Nutzeraktivitäten mittels stationärer und mobiler Hardware zu erkennen.

Ein weiteres Ziel der vorliegenden Arbeit ist, für ausgewählte Wohlbefindensfaktoren Erfassungsverfahren zu entwickeln, die IoT- und Mobile-Computing-Hardware nutzen und die Privatsphäre schützen. Hierbei sollen auch bestehende Forschungsarbeiten aufgegriffen, angepasst und zusätzlich erweitert werden. Bei einigen existenten Erfassungsmethoden sollen Lösungen für die in der Literatur angesprochene Probleme entwickelt werden. Hierzu zählen beispielsweise ein vergleichsweise hoher Nutzeraufwand, störende Systeminteraktionen oder die Auslagerung der Datenverarbeitung auf Systeme von Cloud-Diensten, die zu Risiken beim Schutz der Privatsphäre führen können.

Entwicklung von Assistenz-, Informations- und Empfehlungssystemen zur Verbesserung des Wohlbefindens

Mithilfe spezieller Verfahren zur Erfassung von Wohlbefindensfaktoren können Assistenz-, Informations- und Empfehlungssysteme zur Verbesserung des menschlichen Wohlbefindens entwickelt werden. Entsprechende Systeme nutzen die erfassten Wohlbefindensfaktoren und Nutzereingaben, um Systemausgaben zu erzeugen, die für den Nutzer verständlich sind und ihn unterstützen, sein Wohlbefinden zu verbessern. Die Systeme sollten mittels IoT- und Mobile-Computing-Technologien unaufdringlich in den Alltag von Nutzern integrierbar sein. Hierbei wären zudem die speziellen Anforderungen der Zielgruppe Senioren zu berücksichtigen, da hier bislang Defizite festzustellen sind und Senioren besonders stark von für sie optimierten IT-Systemen profitieren, die ihr Wohlbefinden positiv steigern können. Für die Systementwicklung ist auch die Gewährleistung einer hohen und nachhaltigen Nutzerakzeptanz von großer Bedeutung. Dies sollte durch möglichst einfache Bedienung der Geräte und eine hohe Usability bei den Assistenz-, Informations- und Empfehlungssystemen erreicht werden.

Gewährleistung der Privatsphäre

Insbesondere bei Systemen, die vertrauliche Daten bezüglich Gesundheit und Wohlbefinden verarbeiten, spielt der Schutz der Privatsphäre eine wesentliche Rolle. Bei ubiquitären Systemen tritt zudem das Problem auf, dass diese teilweise kaum sichtbar im Hintergrund agieren und hierdurch eine permanente Überwachung von Nutzern möglich wäre. Bei Sprachassistenten, die grundsätzlich permanent Sprachaufzeichnungen über das Internet an Cloud-Dienste übermitteln können, wurde dies beispielsweise in einer Studie von McLean und Osei-Frimpong (2019) als wesentlicher Grund angegeben, diese nicht zu verwenden. Somit besteht ein Ziel dieser Arbeit auch darin, bereits bei der Entwicklung der Verfahren und Systeme auf den Schutz der Privatsphäre zu achten.

1.4 Struktur der Arbeit

Abbildung 1.1 zeigt in einem Überblick die Struktur dieser Arbeit auf. Teil II dieser Arbeit beinhaltet konzeptuelle und technische Grundlagen, die für die Realisierung der Forschungsziele benötigt wurden. In Teil III werden zunächst die fünf Wohlbefindenskategorien behandelt. Diese Kapitel enthalten jeweils Motivationen, Möglichkeiten der Erfassung entsprechender Wohlbefindensfaktoren, Probleme die mittels eigener Untersuchungen gelöst wurden und die Präsentation der Untersuchungen. Das abschließende Kapitel von Teil III behandelt das CARE-Projekt und darauf basierende Untersuchungen. Hierbei wurden alle fünf Wohlbefindenskategorien, in unterschiedlichem Umfang, in einem kontextsensitiven System für Senioren eingesetzt. In Teil IV werden die einzelnen Teile dieser Arbeit und wissenschaftliche Beiträge in einer Zusammenfassung dargestellt. Dieser Teil enthält zudem den Anhang, indem unter anderem ein Überblick über die wissenschaftlichen Publikationen mit eigener Beteiligung gegeben wird.

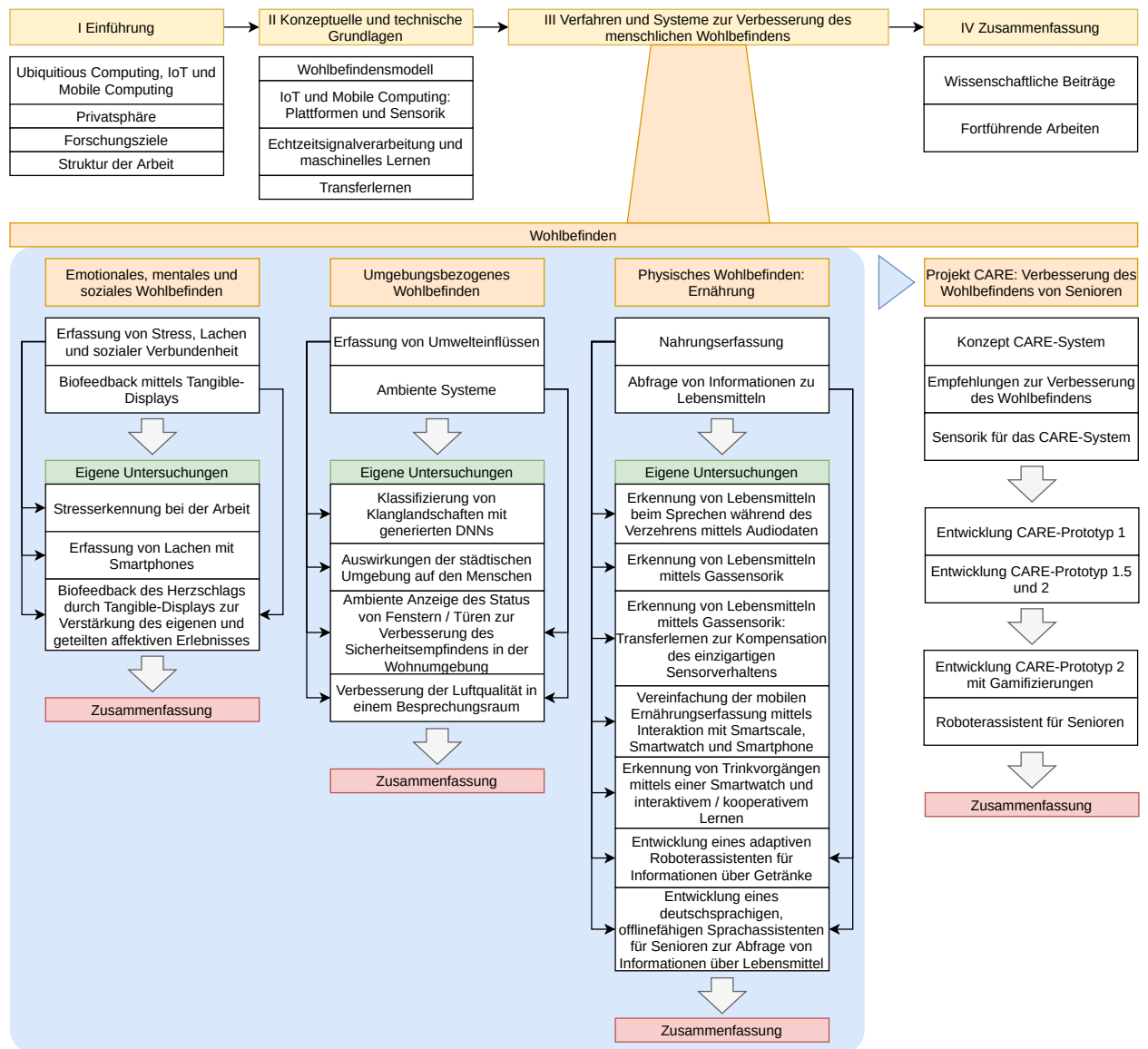


ABBILDUNG 1.1: Struktur der Arbeit.

Teil II

Konzeptuelle und technische Grundlagen

2	Wohlbefindensmodell	13
3	IoT und Mobile Computing: Plattformen und Sensorik	17
3.1	Mikrocontroller und Einplatinencomputer	17
3.2	Mobile Geräte	19
3.3	Smarthome	23
3.4	Verwendete Sensorik	24
3.5	Überblick Gassensorik	30
4	Echtzeitsignalverarbeitung und maschinelles Lernen	39
4.1	Kontinuierliche Daten	39
4.2	Diskrete / eventbasierte Daten	41
4.3	Maschinelles Lernen	45
5	Transferlernen	49
5.1	Definitionen für Domäne, Aufgabe und Transferlernen	50
5.2	Probleme beim Transferlernen	50
5.3	Kategorien von Transferlernen	51
5.4	Ansätze beim Transferlernen	53
5.5	Transferlernen beim Deep Learning	53

2 Wohlbefindensmodell

Die Begriffe Gesundheit und Wohlbefinden bzw. Wohlergehen sind nicht einheitlich definiert, besitzen allerdings einen direkten Bezug zueinander und können sich zudem gegenseitig beeinflussen. Folgende bekannte Definition von Gesundheit wird von der World Health Organization (WHO) seit 1946 verwendet: „Gesundheit ist ein Zustand des vollständigen körperlichen, geistigen und sozialen Wohlergehens und nicht nur das Fehlen von Krankheit oder Gebrechen.“ (siehe deutsche Übersetzung der Verfassung der Weltgesundheitsorganisation von 1946¹). Holdsworth (2019) differenzierte die Begriffe Gesundheit, Wellness und Wohlbefinden mittels verschiedener Literaturquellen. Gesundheit ist hierbei in ihrer traditionellen Form die Basis, bei der das Fehlen von Krankheit Voraussetzung ist. Bereits die Definition der WHO geht allerdings darüber hinaus. Der Begriff „Wellness“ bezieht neben dem Fehlen von Krankheiten einen Lebensstil (engl. lifestyle) zur Verhinderung von Krankheiten ein. Im Begriff Wohlbefinden ist Wellness eingeschlossen, es wird jedoch zusätzlich Freude berücksichtigt.

Dies ist allerdings nur eine mögliche Eingrenzung der Begriffe. Da für das Wohlbefinden keine einheitlichen Definitionen existieren und es für verschiedene Zwecke notwendig ist, quantitative Werte für Vergleiche zu bestimmen, existieren diverse Modelle und zugehörige Fragebögen, um das persönliche Wohlbefinden von Personen zu bewerten. Ein Vergleich dieser Modelle geht über den Rahmen dieser Arbeit hinaus. Entsprechend wurde hier ein pragmatischer Ansatz gewählt, der im Zuge des CARE-Projekts (siehe Kapitel 9) genutzt und in der eigenen Arbeit (Rist, Seiderer, Hammer, Mayr & André, 2015) publiziert wurde. Er basiert auf zwei existierenden Frameworks: dem „National Accounts of Well-being Framework“ der *new economic foundation* (NEF) (Michaelson et al., 2009) sowie dem *SCL/PRB index of well-being in older population* vom *Stanford Center on Longevity* (SCL) und *Population Reference Bureau* (PRB) (Kaneda et al., 2011). Das Ziel des Frameworks der NEF war es, das subjektive Wohlbefinden von Bürgern aus verschiedenen EU-Ländern zu messen. Als Hauptindikatoren dienen das persönliche sowie soziale Wohlbefinden, die in weitere Unterkategorien unterteilt werden. Die tiefsten Ebenen dieser hierarchischen Unterteilung sind mit konkreten Bestandteilen eines Fragebogens verknüpft, die für die Messung in Studien zum Wohlbefinden genutzt werden können. Der SCL/PRB-Index bestimmt Wohlbefinden in vier Domänen: materielles, physisches, kognitives sowie soziales Engagement und emotionales Wohlbefinden. Die Indikatoren werden verwendet, um eine Punktzahl in jeder Domäne vergeben zu können.

In dieser Arbeit wird, wie beim Framework der NEF, ein hierarchisches Modell der Wohlbefindensfaktoren genutzt. Es wurde allerdings dahingehend vereinfacht, dass nur Kategorien enthalten sind, die in Assistenzsystemen wie beispielsweise CARE, eingesetzt werden können. Beim CARE-System handelte es sich um ein personalisiertes, kontextsensitives Empfehlungssystem zur Verbesserung des Wohlbefindens von Senioren. Für ein derartiges System spielt z. B. das Einkommen des Nutzers keine relevante Rolle; für die Bestimmung des materiellen Wohlbefindens wäre dies jedoch ein wichtiger Faktor. Das hier eingesetzte Modell besteht daher aus folgenden Hauptkategorien: emotionales, mentales, soziales, umgebungsbezogenes (engl. environmental) und physisches Wohlbefinden (siehe Abbildung 2.1). Die Kategorien sind vom SCL/PRB-Index inspiriert, wobei anstelle von kognitivem Wohlbefinden der Begriff „mentales Wohlbefinden“ verwendet wird. Zudem ist das Modell um die Kategorie umgebungsbezogenes Wohlbefinden ergänzt worden. Im Rahmen des CARE-Projekts wurde die Plausibilität des Modells in strukturierten

¹Datei von der Webseite des Bundesrates der Schweiz <https://www.admin.ch/opc/de/classified-compilation/19460131/201405080000/0.810.1.pdf> (22.01.2020)

Interviews mit Senioren überprüft. Eine Zusammenfassung enthält Abschnitt 9.2.1. Grundsätzlich besteht durch die Berücksichtigung der Umgebung eine große Ähnlichkeit zu den vier Domänen des validierten WHOQOL-BREF-Lebensqualität-Fragebogens der WHO (Skevington et al., 2004): „physisch“, „psychologisch“, „soziale Beziehungen“ und „Umwelt“. Dieser stellt eine von 100 auf 26 Fragen verkürzte Version des WHOQOL-100-Fragebogens dar. Beide Fragebögen sind etabliert und wurden bereits in einer großen Anzahl an Studien eingesetzt.

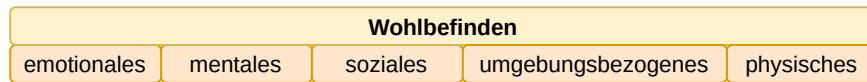


ABBILDUNG 2.1: Überblick über das Wohlbefindensmodell.

Im Nachfolgenden werden die einzelnen Kategorien mit Beispielen erläutert. Hierbei werden zudem beispielhaft Empfehlungen des CARE-Systems angegeben, die potenziell eine positive Auswirkung auf das Wohlbefinden besitzen konnten.

Emotionales Wohlbefinden Zur Bestimmung des emotionalen Wohlbefindens wird beim National-Accounts-of-Well-being-Framework von Michaelson et al. (2009) das Gleichgewicht zwischen der Häufigkeit des Erfahrens positiver und negativer Emotionen herangezogen. Dementsprechend können bei der Erfassung durch Systeme Indikatoren für positive sowie negative Emotionen berücksichtigt werden. Beispielsweise ist „Lachen“ meistens ein Ausdruck positiver Emotion, dahingegen kann das Auftreten von chronischem Stress zu negativen Emotionen führen. Beim CARE-System wurden z. B. Lachübungen und Witze angeboten, um das Gleichgewicht zugunsten positiver Emotionen zu verschieben.

Mentales Wohlbefinden Positives mentales Wohlbefinden wird im Bericht der NEF in Kurzform als sich „gut fühlen und gut funktionieren“ zusammengefasst (Aked et al., 2008). Das emotionale Wohlbefinden ist hierbei eng mit dem Mentalen verknüpft. Ein wichtiger negativer Faktor, der auch von Systemen erfasst werden kann, ist Stress. Dieser tritt beispielsweise bei einer mentalen Überforderung bei der Arbeit auf, wodurch die Produktivität und Qualität der Arbeit gemindert werden. Bei CARE wurde zum mentalen Wohlbefinden auch die Kognition gezählt, da diese u. a. mit der Erinnerungsfähigkeit gemessen werden kann. Entsprechend waren im CARE-System Übungsempfehlungen enthalten, die beispielsweise das Gedächtnis trainieren sollten.

Soziales Wohlbefinden Zum sozialen Wohlbefinden zählen gemäß Michaelson et al. (2009) u. a. enge, positive Beziehungen zur Familie, zu Freunden und Verwandten. Als Indikatoren sind daher beispielsweise die Erfassung positiver Unterhaltungen, bei denen viel gelacht wird oder die Erkennung von Gruppenaktivitäten, wie das gemeinsame Ansehen von Filmen, geeignet. Diesbezügliche Empfehlungen mit der zusätzlichen Möglichkeit einer Telekommunikation mittels virtueller Postkarten waren daher Bestandteil des CARE-Systems.

Umgebungsbezogenes Wohlbefinden Mit umgebungsbezogenem Wohlbefinden werden in dieser Arbeit alle Umgebungsfaktoren kategorisiert, die menschliches Wohlbefinden beeinflussen können. Hierzu zählen beispielsweise das Raumklima in der Wohnung und lokale Klimazonen bei einem Aufenthalt in einer Stadtumgebung. Diese Umweltparameter können mit entsprechender mobiler und stationärer Sensorik von Systemen erfasst werden. Vom CARE-System wurde z. B. eine Empfehlung zum Lüften des Raumes bei entsprechend schlechter Raumluft gegeben.

Physisches Wohlbefinden Physische Aktivitäten werden vom National-Accounts-of-Well-being-Framework sowie SCL/PRB-Index berücksichtigt, da dies Rückschlüsse über die körperliche

Verfassung von Menschen ermöglicht, die direkten Einfluss auf deren Wohlbefinden besitzt. Indikatoren hierfür liefern beispielsweise Bewegungsdaten tragbarer Geräte. Vom CARE-System wurden als Empfehlungen z. B. Spaziergehen oder das Durchführen einer vom System überwachten, leichten Hantelübung vorgeschlagen.

Bei der Evaluation des Wohlbefindensmodells im Rahmen des CARE-Projekts wurde gesunde Ernährung als wesentlicher Bestandteil des physischen Wohlbefindens von den Studienteilnehmern betrachtet (Rist et al., 2015). Auch aus medizinischer Sicht besitzt die Ernährung einen äußerst wichtigen Einfluss auf den menschlichen Körper (Physis). Daher wurde Ernährung als Teil des physischen Wohlbefindens betrachtet. Gesunde Ernährung kann beispielsweise von einem System mit einem digitalen Ernährungstagebuch überwacht werden. Da z. B. eine hinreichende tägliche Flüssigkeitsaufnahme insbesondere bei Senioren problematisch sein kann, gab das CARE-System entsprechende Empfehlungen.

3 IoT und Mobile Computing: Plattformen und Sensorik

Dieses Kapitel beinhaltet einen Überblick zu stationärer und mobiler Sensorik sowie IoT-Geräten. Insbesondere werden Sensoren und Geräte genauer betrachtet, die in dieser Arbeit verwendet wurden. Ein aktueller, ergänzender Überblick über IoT-Technologien ist in der Arbeit von Sultanow und Chircu (2019) verfügbar.

3.1 Mikrocontroller und Einplatinencomputer

In diesem Abschnitt werden alle Hardware-Plattformen vorgestellt, die in dieser Arbeit für mobile oder stationäre Prototypen verwendet wurden. Dadurch, dass es diverse Mikrocontroller und Einplatinencomputer gibt, mussten diese bezüglich ihrer Performanz, Konnektivität, Größe und Stromverbrauch für Projekte in dieser Arbeit ausgewählt werden. Eine bedeutsame Rolle spielte hierbei auch die Möglichkeit einer zeiteffizienten Softwareentwicklung, die insbesondere bei der Arduino- und Raspberry-Plattform durch die Verfügbarkeit diverser Tutorials und Programmbibliotheken vorhanden ist. In Tabelle 3.1 ist eine Auswahl relevanter, technischer Spezifikationen angegeben. In den Abbildungen 3.1 und 3.2 sind Bilder (nicht maßstabsgetreu) der Platinen dargestellt.

Name	Typ	SoC / Microcontrol.	CPU-Kerne	Takt (MHz)	Logik-pegel (V)	Ram	Speicher (KiB)	I ² C	USB	Blue-tooth	WLAN	LAN	OS
Raspberry Pi 3 B	SBC	BCM2837	4	1200	3,3	1024 MiB	μSD	ja	ja	ja	ja	ja	ja
Raspberry Pi 3 B+	SBC	BCM2837B0	4	1400	3,3	1024 MiB	μSD	ja	ja	ja	ja	ja	ja
Raspberry Pi Zero W	SBC	BCM2835	1	1000	3,3	512 MiB	μSD	ja	ja	ja	ja	nein	ja
(Arduino) Uno	μC	ATmega328P	1	16	5	2 KiB	32	ja	nein	nein	nein	nein	nein
(Arduino) Nano	μC	ATmega328P	1	16	5	2 KiB	32	ja	nein	nein	nein	nein	nein
(Arduino) (Pro) Micro	μC	ATmega32U4	1	8 / 16	3,3 / 5	2,5 KiB	32	ja	nein	nein	nein	nein	nein
(Arduino) Pro Mini	μC	ATmega328P	1	8 / 16	3,3 / 5	2 KiB	32	ja	nein	nein	nein	nein	nein
D1 mini	μC	ESP8266	1	80	3,3	96 KiB	4096	ja	nein	nein	ja	nein	ja
nRF51822 Breakout	μC	nRF51822	1	16	3,3	32 KiB	256	ja	nein	ja	nein	nein	nein

TABELLE 3.1: Übersicht über die verwendeten Einplatinencomputer (single board computer = SBC) und Mikrocontroller (μC). Quellen: ¹

Schnittstellen Die leistungsfähigsten Geräte (CPU-Leistung, Arbeitsspeicher), die zudem fähig sind, mit einem Linux-Betriebssystem genutzt zu werden, sind hierbei die Einplatinencomputer. Sie besitzen jedoch auch den höchsten Stromverbrauch. Alle eingesetzten Raspberry Pis verfügten mindestens über eine Bluetooth- und WLAN-Schnittstelle. Nur die vergleichsweise leistungsfähigen

¹(17.06.2020): <https://www.raspberrypi.org/products/raspberry-pi-3-model-b/>
<https://www.raspberrypi.org/products/raspberry-pi-3-model-b-plus/>
<https://www.raspberrypi.org/products/raspberry-pi-zero-w/>
<https://www.microchip.com/wwwproducts/en/ATmega328>
<https://www.microchip.com/wwwproducts/en/ATmega32U4>
https://www.espressif.com/sites/default/files/documentation/0a-esp8266ex_datasheet_en.pdf
<https://www.nordicsemi.com/Products/Low-power-short-range-wireless/nRF51822>

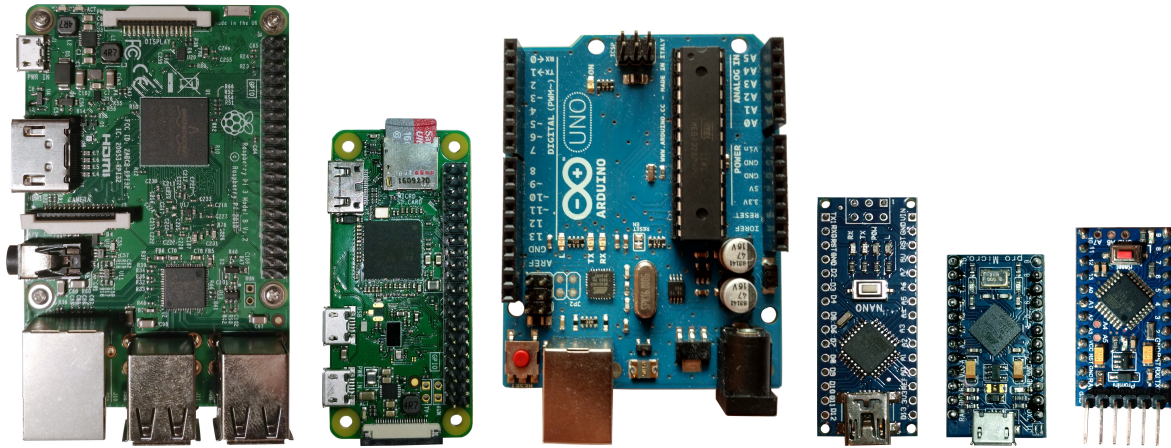


ABBILDUNG 3.1: Von links: Raspberry Pi: 3B / Zero W, Arduino: Uno / Nano (Klon) / Micro (Klon) / Pro Mini (Klon).

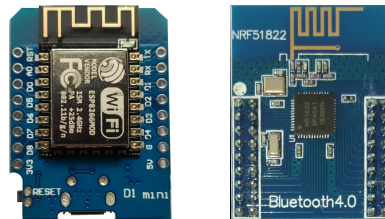


ABBILDUNG 3.2: Von links: D1 mini (ESP8266), NRF51-Breakoutboard.

Mikrocontroller ESP8266 und nRF51822 besitzen ein integriertes WLAN- bzw. Bluetooth-Funkmodul. Der ATmega-Mikrocontroller ATmega32U4 kann direkt per USB-Schnittstelle (Client) angesteuert werden. Beim ATmega328P wird ein zusätzlicher Chip bzw. Mikrocontroller benötigt, der von der USB- auf eine serielle Schnittstelle wandelt. Beim Arduino Nano und Uno sind diese Konverter auf der Platine integriert. Alle aufgelisteten Geräte verfügen über verschiedene Anzahlen an digitalen und analogen Anschlüssen, an die externe Komponenten und Schaltungen angeschlossen werden können. Je nach Chip und teilweise Versorgungsspannung liegt der Logikpegel bei den digitalen Pins bei 3,3 oder 5 V. Falls eine Komponente mit 5 V-Logikpegel beispielsweise an einen ESP8266 mit 3,3 V-Logikpegel angeschlossen werden soll, muss ein entsprechender Pegelwandler eingesetzt werden. Der I²C-Bus kann per Hardware und Software implementiert werden. Es werden hierfür zwei digitale Pins benötigt. Der Bus erlaubt beispielsweise das Ansteuern einer Vielzahl an Sensoren. USB-Host-Funktionalität besitzen in dieser Übersicht nur die Raspberry-Pi-Einplatinencomputer. An diese können daher andere USB-Geräte angeschlossen werden. Der System- / Datenspeicher ist bei diesen Einplatinencomputer mittels einer µSD-Karte umgesetzt, wodurch verschiedene Speichergrößen möglich sind. Auf diesem Speicher befindet sich auch das Linux-Betriebssystem. Bei den Mikrocontrollern sind die Programm- / Datenspeicher im Chip integriert oder ein externer Speicherchip (ESP8266) ist auf der Platine vorhanden. Insbesondere die ATmega328-Chips besitzen nur 32 KiB Speicher für Programme und Daten. Die vollständigen Spezifikationen der einzelnen Chips sind in den entsprechenden Datenblättern ersichtlich.

Software Bei der Implementierung von Prototypen spielt auch die Software-Unterstützung eine erhebliche Rolle. Mit Einplatinencomputern können die meisten Programme genutzt werden, die unter Linux und der ARM-Architektur lauffähig sind. Bei Mikrocontrollern stellt die Arduino-Kompatibilität eine relevante Eigenschaft dar, da hierdurch eine Vielzahl an Programmbibliotheken zur Verfügung steht. Besonders hervorzuheben ist diese Unterstützung für die Chips der ATmega-Reihe, die klassischerweise für die ersten Arduino-kompatiblen-Boards eingesetzt wurden. Mithilfe

der Community werden jedoch auch viele weitere Mikrocontroller zumindest teilweise unterstützt, wodurch (nahezu) identische Programme auf diesen lauffähig sind. Beispielsweise verfügt der ESP8266 über eine gut geeignete Arduino-Unterstützung. Inkompatibilitäten mit bestimmten Programmbibliotheken können allerdings auftreten. Dies gilt besonders für den nRF51822-Chip. In solchen Fällen muss notfalls auf die Nutzung von Arduino-Code verzichtet werden oder die Bibliotheken müssen entsprechend angepasst werden.

Auswahl geeigneter Plattformen Bei der Auswahl der für einen Prototyp am besten geeigneten Plattform spielen verschiedene Faktoren eine Rolle. Zunächst müssen hierfür die Anforderungen festgelegt werden. Falls die Leistungsfähigkeit eines Mikrocontrollers ausreichend ist, können diese wesentlich energieeffizienter und kompakter sein als Einplatinencomputer. Auch der Kaltstart eines Mikrocontrollers ist dadurch, dass hierbei nur ein Programm ausgeführt wird, wesentlich schneller als das Booten eines komplexen Betriebssystems, bei dem eine Vielzahl an Treibern geladen und Zusatzprogramme gestartet werden müssen. Dadurch, dass bei einem Mikrocontroller in der Regel nur ein einzelnes Programm gestartet wird, können auch unvorhersehbare Wechselwirkungen mit anderen Programmen oder dem Betriebssystem nicht auftreten.

Die Auswahl eines Mikrocontrollers anstelle von einem Einplatinencomputer stellt daher eine Optimierung dar, die zulasten der Flexibilität geht, da hierbei größere Restriktionen bezüglich verwendbarer Software, Schnittstellen, CPU-Leistung, Arbeitsspeicher und Datenspeicher auftreten. Die in dieser Arbeit genutzten Hardware-Plattformen wurden aus diesem Grund gemäß ihrem Anwendungszweck möglichst optimal ausgewählt.

3.2 Mobile Geräte

3.2.1 Systeme

In dieser Arbeit wurden verschiedene kommerzielle, mobile Geräte genutzt. Hierzu zählen Smartphones, Smartwatches und Fitnessarmbänder. Smartphones und Smartwatches verwenden häufig leistungsfähige SoCs mit ARM-Architektur, die teilweise auch für Einplatinencomputer verfügbar sind. Fitness-Tracker setzen eher auf die eingeschränkteren Mikrocontroller mit Bluetooth LE wie z. B. die nRF51-Serie, die zudem kostengünstiger sind.

Smartphones In dieser Arbeit wurden folgende Smartphones für verschiedene Anwendungen eingesetzt: Nexus 4 (für Google durch LG hergestellt), Nexus 6P (für Google durch Huawei hergestellt), Samsung Galaxy S3 und S4. Es wurden ausschließlich Geräte mit Android-Betriebssystem genutzt, da diese im Vergleich zu iOS weltweit verbreiteter (Juli 2020: Android: 74,6 %; iOS: 24,82 %²) und offener bezüglich der Installation eigener (unsignierter) Software sind. Android basiert auf Linux und ist im Kern quelloffen³. Die meisten Android-Installationen enthalten allerdings proprietäre Erweiterungen wie z. B. die Google-Dienste und Treiber. Auf kommerziellen Geräten besitzt der Nutzer keine administrativen Rechte (root-Rechte) und kann daher keine Änderungen am System vornehmen. Alternative Android-Systeme wie die Distribution LineageOS⁴, die auf bestimmten Smartphones installiert werden können, sind weniger beschränkt und ermöglichen mehr Kontrolle über das System.

Die Sensorik von Smartphones umfasst in der Regel mindestens ein Accelerometer, Gyroskop, Magnetometer, einen Helligkeitssensor, eine Kamera und ein Mikrofon. Eine Bestimmung der geografischen Position per Satellit ist meistens ebenfalls möglich. Nutzereingaben können mittels GUIs über relativ große Touchscreens oder durch Spracheingaben erfolgen. Visuelle, auditive und

²<https://gs.statcounter.com/os-market-share/mobile/worldwide/> (01.08.2020)

³<https://source.android.com/> (01.08.2020)

⁴<https://lineageos.org/> (01.08.2020)

taktile Ausgaben können vom Smartphone erzeugt werden. Zur Interaktion mit anderen Geräten stehen eine Bluetooth-, WLAN- und häufig eine Mobilfunkverbindung zur Verfügung.

Smartwatches In dieser Arbeit wurden folgende Smartwatches für verschiedene Anwendungen eingesetzt: Asus ZenWatch 2, Motorola Moto 360 gen 1, Samsung Gear S2. Die Smartwatches von Asus und Motorola verwendeten Android Wear als Betriebssystem, die Samsung Smartwatch setzte Tizen⁵ ein. Smartwatches mit Android Wear besitzen den Vorteil, dass Software von Android-Smartphones grundsätzlich auch auf diesen nutzbar ist. Die eingesetzten Smartwatches waren generell unabhängig von einem Smartphone nutzbar. Allerdings ist im Vergleich zu Smartphones die grafische Schnittstelle durch die kleinen Displays wesentlich eingeschränkter und die Hardware weniger leistungsfähig. Durch ihre Positionierung am Handgelenk sind sie jedoch meist schneller erreichbar und öfters im Blickfeld als ein Smartphone. Auch können sie auf Benachrichtigungen taktil, durch Vibration, aufmerksam machen. Wie beim Smartphone besitzen die meisten Smartwatches mindestens jeweils ein Accelerometer, Gyroskop und Magnetometer als Sensorik. Anhand ihrer Position am Armgelenk ermöglicht eine Smartwatch somit beispielsweise die Auswertung von Armbewegungen. In einigen Fällen ist zusätzlich ein Sensor zur Erfassung der Herzrate integriert. Bei der Motorola Moto 360 gen 1 und Samsung Gear S2 wird hierfür ein optischer BVP-Sensor eingesetzt. Alle drei eingesetzten Smartwatches verfügten über ein Mikrofon, das z. B. Spracheingaben ermöglicht. Die meisten Android-Smartwatches besitzen eine WLAN- und Bluetooth-Schnittstelle. Sie können hiermit mittels der Google-API mit dem Smartphone kommunizieren. Die Google-API nutzt automatisch eine WLAN-Verbindung und die Google-Cloud, falls die Smartwatch zu weit vom Smartphone entfernt liegt. In der Nähe des Smartphones wird direkt eine energiesparendere Bluetooth-Verbindung zwischen den Geräten aufgebaut. Die Nutzung von WLANs durch Smartwatches kann mittels entsprechender Einstellungen verhindert werden, wodurch keine Daten an die Google-Cloud übermittelt werden.

Fitnessarmbänder Im Rahmen dieser Arbeit wurde das Microsoft-Band-2-Fitnessarmband⁶ eingesetzt. Fitnessarmbänder sind in den meisten Fällen durch Nutzung leistungsschwächerer Hardware wesentlich eingeschränkter als Smartwatches und können in der Regel vom Nutzer nicht direkt programmiert werden. Dafür sind diese kompakter als Smartwatches und ihre Batterielaufzeit ist länger. Da Fitnessarmbänder für sportliche Aktivitäten ausgelegt werden, besitzen sie neben mindestens einem Accelerometer, das z. B. für die Bestimmung der Schrittzahl eingesetzt wird, in den meisten Fällen einen Sensor zur Ermittlung der Herzrate. Dies ist auch beim Microsoft Band 2 der Fall. Eine Problematik bei Fitnessarmbändern stellt das Abrufen von Sensordaten mit eigenen Anwendungen dar. In den meisten Fällen wird dies von den Herstellern unterbunden und die Daten können teils nur über deren Clouds oder Hersteller-Apps ausgelesen werden. Teilweise werden die Daten zunächst im internen Speicher des Trackers gespeichert, um sie in bestimmten Zeitintervallen per Bluetooth an ein gekoppeltes Smartphone zu übermitteln, das diese Daten auf die Server der Hersteller überträgt. Eine Echtzeiterfassung der Sensordaten ist nur bei wenigen Modellen, wie dem Microsoft Band 2, möglich und muss auf dem Smartphone stattfinden. Generell verursacht eine Echtzeitübertragung der Sensordaten an ein Smartphone jedoch einen hohen Energiebedarf und es ist mit wesentlich verringerten Akkulaufzeiten zu rechnen.

3.2.2 Mobile Datenverarbeitung

Mobile Geräte sind generell technisch schwieriger umzusetzen als stationäre, da eine Stromversorgung über Batterien / Akkus möglichst lange sichergestellt werden sollte und die Geräte nicht zu schwer oder unhandlich groß sein dürfen. Im Vergleich zu stationären Geräten benötigt deren Elektronik besonderen Schutz gegenüber mechanischer Belastung, hoher Feuchtigkeit, Staub und

⁵<https://www.tizen.org/> (01.08.2020)

⁶<https://support.microsoft.com/de-de/help/4000313/band-hardware-band-2-features-and-functions> (03.12.2019)

elektrostatischen Entladungen. Diese Anforderungen schränken daher auch die Auswahl an Sensorik ein. Auf der Softwareseite müssen alle möglichen Energiesparmaßnahmen genutzt werden, wie z. B. eine intelligente Reduktion der Taktrate des Mikrocontrollers oder das temporäre Abschalten ungenutzter Komponenten. Bei der Datenverarbeitung muss zusätzlich auf energieeffiziente Verfahren und Algorithmen zurückgegriffen werden.

Zur Bestimmung des menschlichen Wohlbefindens sind insbesondere Biosignale von Relevanz, die möglichst permanent vom Nutzer erfasst werden sollten, wodurch mobile Geräte für diese Aufgabe prädestiniert sind. Es existieren spezielle mobile Geräte, wie das biosignalsplux⁷, die Biosignale in hoher Qualität aufzeichnen oder per Funk übertragen können. Im Gegensatz hierzu spielt bei kommerziellen Smartphones, Smartwatches und Fitness-Trackern die Datenqualität bei Sensoren nicht die wichtigste Rolle. Relevanter ist für die Hersteller, dass bestimmte Funktionalitäten ermöglicht werden (z. B. Bestimmung der Herzrate), dass die Sensorik kostengünstig, mechanisch geschützt, energie- und platzsparend integriert werden kann. Zudem sollte die Handhabung durch den Nutzer möglichst problemlos und robust sein, sodass beispielsweise keine manuelle Kalibrierung oder das Anbringen von externen Klebeelektroden (z. B. für EKG) notwendig ist. Trotz ihrer häufig schlechteren Datenqualität ist der Einsatz kommerzieller mobiler Geräte für Feldstudien besonders interessant, da es hierbei von großer Bedeutung ist, dass die Datenaufzeichnung weitestgehend unauffällig im Hintergrund durchgeführt werden kann. Ansonsten kann „normales“ Nutzerverhalten nur begrenzt beobachtet werden.

Bestimmung der Herzrate aus dem BVP-Signal

Viele Smartwatches und Fitnessarmbänder verfügen über einen Sensor zur Ermittlung der Herzrate aus dem Blutvolumenpuls-Signal (BVP), das optisch, nichtinvasiv erfasst werden kann. In mehreren eigenen Projekten wurde die Herzrate sowie teilweise die Herzratenvariabilität aus dem BVP-Signal ermittelt: siehe Abschnitt 6.3.1, 6.3.3 und 7.3.2. Im Folgenden wird zunächst das Prinzip der Erfassung des BVP-Signals erörtert. Anschließend wird dessen Verarbeitung anhand des Sensors, der auch in Abschnitt 6.3.3 für den PiHeart-Prototyp mit einem akkubetriebenen, tragbaren Raspberry Pi Zero W eingesetzt wurde, beschrieben. Es werden hierbei zwei Auswertungsmethoden beispielhaft erklärt, die teilweise auch bei kommerziellen Geräten eingesetzt werden. Hierzu ergänzende Grundlagen und weiterführende Verfahren zur Ermittlung des BVP-Signals mittels tragbarer Sensorik sind in der Arbeit von Tamura et al. (2014) enthalten.

Messprinzip Bei jedem Herzschlag wird Blut durch die Adern gepumpt. Hierbei wechseln sich die Phasen „Systole“ (Blut strömt aus dem sich zusammenziehenden Herzen) und „Diastole“ (Blut strömt in das Herz) ab. Durch die Änderung des Drucks beim Pumpen verändert sich das Volumen der Blutgefäße. Die größten Änderungen entstehen hierbei in der Arterie, bei der während der Systole ein größeres Blutvolumen enthalten ist, als in der Diastole (Tamura et al., 2014). Diese Änderungen können mit einem optischen Sensor auf der Haut erfasst werden, da sich hierbei die Reflexion einer konstanten Lichtquelle (meist rot oder grün) messbar verändert.

Erfassung des BVP-Signals In diesem Abschnitt wurde der MAX30102-BVP-Sensor des PiHeart-Prototyps eingesetzt, der so präpariert war, dass er stabil an der Fingerkuppe platziert werden konnte (siehe Abbildung 3.3 rechts). Generell wären auch andere Positionen am Körper möglich, wie z. B. am Ohr läppchen. Messungen am Handgelenk sind problematischer, da hier die Änderungen des Volumens geringer ausfallen und je nach Position zusätzlich Körperbehaarung eine Rolle spielen kann. Der hier verwendete MAX30102-Sensor zeigte bei Versuchen am Arm und Handgelenk mit verschiedenen Einstellungen (Empfindlichkeit, Helligkeit der LEDs) und Abständen keine akzeptable Signalqualität. Eine möglichst stabile Position des Sensors sollte gewährleistet

⁷<https://www.biosignalsplux.com/> (02.08.2020)

werden, da bereits kleine Bewegungen Artefakte im BVP-Signal erzeugen, was zu Messfehlern führt. Dies gilt auch für kommerzielle Produkte, die teilweise das Signal vollständig verwerfen oder als ungenau markieren (z. B. Microsoft Band 2), wenn diese stärkere Bewegungen mittels ihrer Beschleunigungssensoren registrieren. Fremdlichtquellen können das Signal ebenfalls stören, dementsprechend nutzt der MAX30102-Sensor eine spezielle integrierte Filtertechnik. Es wurden in dieser Arbeit zur Berechnung der Herzrate zwei Ansätze verfolgt. Zum einen können die zeitlichen Abstände der maximalen Signalausschläge (Peaks) im BVP-Signal genutzt werden, zum anderen ist es möglich, die Herzrate per Fouriertransformation mittels des FFT-Algorithmus im Spektrum zu bestimmen.

Herzrate: Signalausschläge (Peaks) Die Berechnung der Abstände zwischen den Schlägen entspricht dem Biosignal „Inter-Beat-Interval“ (IBI). Dieses kann in die Herzrate umgerechnet werden. Ein Abstand von 0,8 Sekunden entspricht beispielsweise 75 (60 sec / 0,8 sec) Schlägen pro Minute (engl. beats per minute: bpm). Das entwickelte Programm erfasste zunächst Rohdaten des BVP-Sensors mit einer Abtastrate von 100 Hz, hierbei wurde der FIFO-Puffer (32 Samples) des Sensors ausgenutzt. Das Signal wurde danach in 5-Sekunden-Fenster unterteilt, wodurch immer mehrere Herzschläge erfasst wurden. Daraufhin wurde ein Butterworth-Bandpassfilter der 5. Ordnung mit einer unteren Grenzfrequenz von 1 Hz (60 bpm) und einer oberen Grenzfrequenz von 3,5 Hz (210 bpm) angewandt. Um eine Phasenverschiebung zu verhindern, wurde der *filtfilt*-Befehl von *scipy* genutzt, der den Butterworth-Filter einmal vorwärts und rückwärts anwendet⁸. Eine Filterung des Signals ist notwendig, falls Artefakte durch Bewegungen vorhanden sind. Sonst ist die Signalqualität für eine Weiterverarbeitung bereits ausreichend hoch (siehe Abbildung 3.3 links).

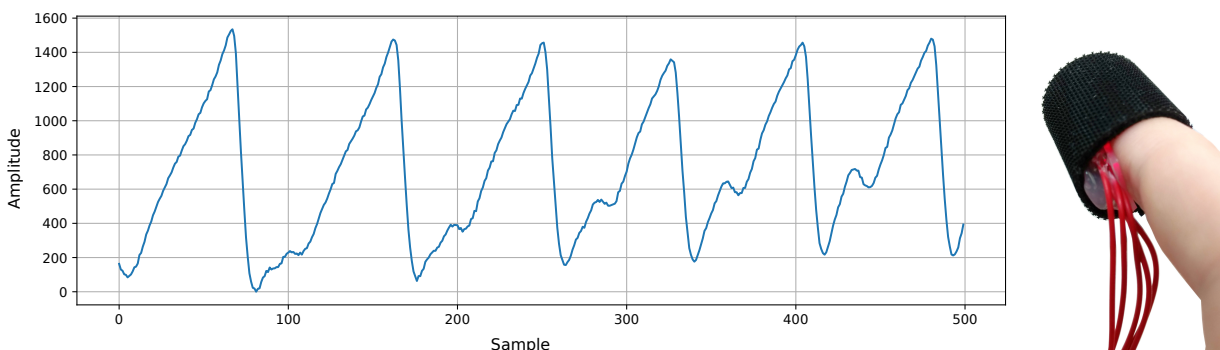


ABBILDUNG 3.3: 5 Sekunden unverarbeitetes Rohsignal (Abtastrate: 100 Hz) mit sehr guter Qualität des MAX30102-BVP-Sensors an der Fingerkuppe (siehe rechts).

Als nächster Schritt wurde ein Algorithmus angewandt, der die Maxima findet (Peakdetector). Hierfür existieren ebenfalls verschiedene Verfahren, wofür beispielsweise auch eines in (mobile)SSI integriert ist, das allerdings für den MAX30102-Sensor nicht hinreichend exakt funktionierte. In diesem Fall wurde das Python-Modul *PeakUtils*⁹ genutzt. Zunächst wurde die Baseline vom Signal abgezogen, die mit einer iterativen polynomialen Regression berechnet wurde und somit Grundschwankungen in der Amplitude reduzierte, die beispielsweise bei Änderungen des Sensor-Anpressdrucks auftreten können. Die Anzahl der Iterationen wurde auf zwei festgelegt. Auf diese verarbeiteten Daten wurde anschließend die Hauptfunktion des *PeakUtils*-Moduls angewandt, um die Positionen der Peaks zu bestimmen. Hierbei kann ein Mindestabstand angegeben werden, der verhindert, dass Signalausschläge mit zu geringer Zeitspanne zueinander ermittelt werden. Da die Sensorik nicht während hoher körperlicher Belastung genutzt wurde, wurden 200 bpm als kürzester Zeitabstand angegeben, was einem sehr hohen Puls entspricht. Das Programm errechnete anschließend die zeitlichen Abstände zwischen den einzelnen Signalausschlägen und mittelte diese.

⁸<https://docs.scipy.org/doc/scipy-0.18.1/reference/generated/scipy.signal.filtfilt.html> (12.03.2020)

⁹<http://peakutils.readthedocs.io/en/latest/> (12.03.2020)

Daraufhin wurde aus dem Mittelwert die Herzrate berechnet. Bei Tests auf dem Raspberry Pi Zero W stellte sich heraus, dass die hier eingesetzte Echtzeitberechnung der Peak-Positionen zu viel Zeit benötigte. Weitere Optimierungen wären bei dieser Hardware und diesem Sensor notwendig, die auch die Stabilität der Erkennung einzelner Schläge reduzieren können. Falls einzelne Herzschläge nicht bestimmt werden müssen, ist die nachfolgende Methode besser geeignet.

Herzrate: Fouriertransformation Eine weitere Methode zur Bestimmung der Herzrate ist die Ermittlung der Herzschlagfrequenz aus einem Frequenzspektrum, das mittels der Fouriertransformation aus einem Zeitintervall der Messung berechnet wurde. Die Fouriertransformation ist ein essenzieller Bestandteil diverser Signalverarbeitungsalgorithmen und wurde für viele CPU-Architekturen hinsichtlich Rechengeschwindigkeit deutlich optimiert. Diese Fast-Fourier-Transformation (FFT) ist hinreichend schnell, sodass sie auch bei weniger leistungsfähigen Computern, wie dem Raspberry Pi Zero W, Anwendung finden kann. Der implementierte Algorithmus erfasste zunächst analog zum Peak-Erkennungsverfahren die BVP-Daten aus dem FIFO-Puffer des MAX30102 mit einer Abtastrate von 100 Hz. Die Fenstergröße für die sogenannte Short-Time-Fouriertransformation (STFT) wurde auf 30 Sekunden (3000 Samples) festgelegt, worüber die FFT ausgeführt wurde. Es wurden hierbei zu 75 % (2250 Samples) überlappende Fenster genutzt, um alle 7,5 Sekunden die Herzrate zurückgeben zu können, nachdem am Anfang ein Fenster nach 30 Sekunden vollständig gefüllt wurde. Die Samples eines Fensters wurden vor der FFT normalisiert, indem alle Samples durch den Maximalwert der Datenwerte im Fenster geteilt wurden. Daraufhin wurde die FFT angewandt, die eine Frequenzauflösung von ca. 0,033 Hz (100 Hz/3000) (ca. 2,0 bpm) in einem Frequenzbereich von 0 bis 50 Hz lieferte. Für die Berechnung der Herzrate wurden Frequenzen außerhalb von ca. 0,6667 Hz und 5 Hz (40 und 300 bpm) aus dem Spektrum entfernt (gefiltert). Aus dem gefilterten Spektrum wurde anschließend der Index des Maximalwerts der Absolutwerte des Real-Teils der Imaginären-Zahlen berechnet und die Frequenz bestimmt, die am stärksten im BVP-Signal vorhanden war. Diese entspricht bei einem sauberen Signal der Herzrate.

Diese frequenzbasierte Herzratenberechnung arbeitete auf dem Raspberry Pi Zero W deutlich stabiler als die Berechnung mittels Peaks im zeitlichen Bereich.

3.3 Smarthome

In einem Smarthome wird eine Vielzahl an Sensoren verbaut, die dazu dienen Aktuatoren automatisiert anzusteuern oder auch Informationen für die Bewohner aufbereitet zur Verfügung zu stellen. Es existieren derzeit mehrere Systeme für den Smarthome-Bereich. Im Folgenden werden insbesondere funkbasierte Systeme behandelt, da diese bei Gebäuden leicht nachrüstbar sind und auch für Feldstudien temporär installiert werden können.

Funkübertragung

Folgende Funkprotokolle sind in Deutschland derzeit verbreitet und werden in Smarthome-Systemen verschiedener Hersteller in ihren Produkten eingesetzt:

- EnOcean (868 MHz - EU)
- Homematic (IP) (868 MHz)
- Z-Wave (Plus) (868 MHz - EU)
- DECT ULE (1,9 GHz)
- Bluetooth (LE) (2,4 GHz)
- WiFi (2,4 GHz)
- ZigBee (2,4 GHz)

Zwischen den verschiedenen Funkprotokollen besteht keine Interkompatibilität. Hardware für das Homematic-Protokoll wird ausschließlich vom Hersteller eQ-3 entwickelt und produziert. Für das ZigBee-Protokoll werden beispielsweise die Philips-Hue-Produkte angeboten.

Die verschiedenen Protokolle besitzen unterschiedliche Vor- und Nachteile. Generell sind Funkprotokolle, die in einem höheren Frequenzbereich (z. B. 2,4 GHz) genutzt werden, schlechter für höhere Reichweiten geeignet, da diese stark durch Wände, Menschen oder sonstige Hindernisse in einem Gebäude gedämpft werden und im Vergleich zu niedrigeren Frequenzen mehr Energie benötigt wird, um eine vergleichbare Reichweite zu erreichen. Die größere Bandbreite, die bei höheren Frequenzen verfügbar ist, ist zum Senden von Sensordaten wie der Temperatur oder Empfangen von Befehlen nicht notwendig.

Der Stromverbrauch spielt bei batteriebetriebenen Sensoren und Aktuatoren eine große Rolle. Das Funkprotokoll sollte dementsprechend möglichst effizient sein. Einen besonderen Weg gehen EnOcean-Sensoren, die weitestgehend batterieless arbeiten und mithilfe z. B. von Solarzellen sogenanntes „Energy Harvesting“ betreiben, um die Elektronik mit Strom zu versorgen. Diese Form der Energiegewinnung ist jedoch nicht für jede Art von Sensorik ausreichend.

WiFi ist wegen des vergleichbar hohen Stromverbrauchs vor allem für an das Stromnetz angebundene Sensoren oder Aktuatoren interessant. Geräte können sich hiermit direkt mit dem lokalen Netzwerk und Internet verbinden und sind daher von überall erreichbar. Eine direkte Ansteuerung mit anderen Netzwerkgeräten, wie beispielsweise einem Smartphone, ist möglich. Bluetooth-Geräte können ebenfalls direkt vom Smartphone und direkt verbundenen anderen Geräten angesteuert werden. In der Low-Energy-Variante (LE) ist die Reichweite reduziert, jedoch kann ein akzeptabler Stromverbrauch erreicht werden. Für andere Funkprotokolle wird eine sogenannte Bridge im lokalen Netzwerk benötigt, die zwischen dem kabelgebundenen oder kabellosen Netzwerk und dem entsprechenden Protokoll die Daten austauscht.

Hinsichtlich der Netzwerkarchitektur sind viele ältere Protokolle sternförmig aufgebaut. Es gibt einen zentralen Transceiver, über den alle Nachrichten geleitet werden müssen. Die Reichweite kann nur mittels speziellen Repeatern erweitert werden. Fast alle der zuvor aufgelisteten Protokolle erlauben allerdings in neueren Ausführungen den Aufbau von Mesh-Netzwerken, wodurch andere Funknetzwerk-Teilnehmer fähig sind, Daten direkt untereinander auszutauschen und Sensordaten und Befehle weiterzuleiten. Dies ist insbesondere für über eine größere Fläche verteilte Systeme interessant, bei denen nicht alle Teilnehmer die zentrale Einheit erreichen können. Es führt jedoch auch zu höheren Latenzen, je nachdem über wie viele Teilnehmer die Daten weitergeleitet werden müssen.

Plattformen

Wie im vorherigen Abschnitt angesprochen, existiert eine große Anzahl inkompatibler Funkstandards. Falls mehrere proprietäre Standards vereint werden sollen, muss eine Hardware-Plattform für jedes Funkprotokoll ein entsprechendes Funkmodul oder Zugriff auf eine Schnittstelle im lokalen Netzwerk besitzen. Zudem wird eine universelle Hausautomations-Software benötigt. Als nahezu Standard-Plattform haben sich die kostengünstigen und weitverbreiteten Raspberry-Pi-Einplatinencomputer positioniert, wofür die meisten Hersteller zertifizierte Funkmodule und Software anbieten. Die Funkmodule werden hierbei per GPIO- oder USB-Verbindung angeschlossen. Raspberry Pis ermöglichen es zudem vergleichsweise problemlos, eigene Sensoren und Aktuatoren anzubinden, sodass diese als vollständige Smarthome-Zentrale dienen können.

3.4 Verwendete Sensorik

In den folgenden Abschnitten werden ausgewählte Sensoren vorgestellt, die zum Großteil in dieser Arbeit eingesetzt wurden. Gassensorik wird in Abschnitt 3.5 ausführlicher behandelt.

Bewegungssensorik

Ein klassischer Sensortyp im Smarthome ist die Erkennung von Nutzerpräsenz mittels Bewegungsmeldern. Hierfür werden im allgemeinen PIR-Sensoren (Passiv-Infrarot-Sensor), die auf dem pyroelektrischen Effekt basieren, eingesetzt (Hering, Schönfelder et al., 2012, S. 72 f.). Bei Temperaturänderungen verändert sich die elektrische Polarisation des pyroelektrischen Materials wodurch eine messbare Spannung entsteht (Hering, Schönfelder et al., 2012, S. 69 ff.). Die Empfindlichkeit (Verstärkung) ist bei den meisten Bewegungsmeldern justierbar und beeinflusst, neben der Nutzung von Fresnellinsen, den Erfassungsbereich des Sensors. Ein kostengünstiger PIR-Sensor, der in eigenen Schaltungen integriert werden kann, ist in Abbildung 3.4 dargestellt.

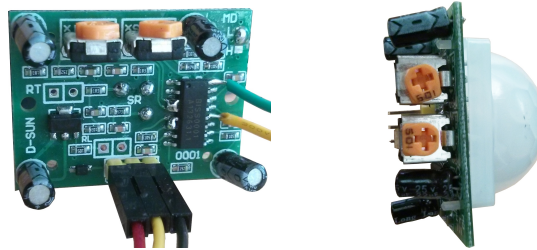


ABBILDUNG 3.4: Einfacher PIR-Sensor mit Auswertungselektronik.

Helligkeitssensorik

In vielen Bewegungsmeldern wie z. B. beim Homematic-System sind Helligkeitssensoren integriert, da diese häufig zum Schalten von Licht in der Dunkelheit eingesetzt werden. Die einfachste Form stellt hierbei ein Fotoresistor dar, der abhängig von der Beleuchtungsstärke den Widerstand verändert. In dieser Arbeit wurden in einigen Projekten auch digitale BH1750-Sensoren¹⁰ eingesetzt, welche die kalibrierte Beleuchtungsstärke in Lux ausgeben. Neben einigen TV- und PC-Bildschirmen besitzen auch die meisten mobilen Geräte Helligkeitssensoren, um deren Bildschirmhelligkeit an das Umgebungslicht dynamisch anzupassen.

Raumklima- und Luftqualitätssensorik

Das Raumklima kann mittels verschiedener Sensoren erfasst werden. Hierzu zählen klassischerweise Temperatur- und Luftfeuchtigkeitssensoren. Zusätzlich werden immer häufiger Feinstaubsensoren im Smarthome eingesetzt.

Temperatur und relative Luftfeuchtigkeit

Für die Messung von Temperatur und relativer Luftfeuchtigkeit werden beispielsweise folgende, kompakte Sensoren eingesetzt: DHT22¹¹ (Temperatur + Luftfeuchtigkeit), Bosch BME280¹² (Temperatur + Luftfeuchtigkeit + Luftdruck) und Sensirion SHT75¹³ (Temperatur + Luftfeuchtigkeit). Die Sensoren sind in Abbildung 3.5 dargestellt. Die Sensoren unterscheiden sich u. a. bezüglich ihrer Messgenauigkeiten, Abmessungen und Preis. In dieser Auswahl ist der SHT75-Sensor der präziseste Sensor (Temperatur: $\pm 0,3^{\circ}\text{C}$; relative Luftfeuchtigkeit: $\pm 1,8\%$), der DHT22 ist der günstigste Sensor und der BME280 der kompakteste (SMD-Bauteil: $2,5 \times 2,5 \times 0,93\text{ mm}$), der zusätzlich den Luftdruck messen kann. Der BME280-Sensor wird als mikroelektromechanisches System (MEMS)

¹⁰https://www.mouser.de/datasheet/2/348/Rohm_11162017_ROHMS34826-1-1279292.pdf (22.02.2021)

¹¹<https://www.sparkfun.com/datasheets/Sensors/Temperature/DHT22.pdf> (19.02.2021)

¹²<https://www.bosch-sensortec.com/media/boschsensortec/downloads/datasheets/bst-bme280-ds002.pdf> (19.02.2021)

¹³https://www.sensirion.com/fileadmin/user_upload/customers/sensirion/Dokumente/2_Humidity_Sensors/Datasheets/Sensirion_Humidity_Sensors_SHT7x_Datasheet.pdf (19.02.2021)

gefertigt und ist u. a. durch seine Kompaktheit und seinen geringen Stromverbrauch zusätzlich für mobile Geräte wie Wearables und Smartphones vorgesehen, wo er auch zur Höhenbestimmung per Luftdruck eingesetzt werden kann.

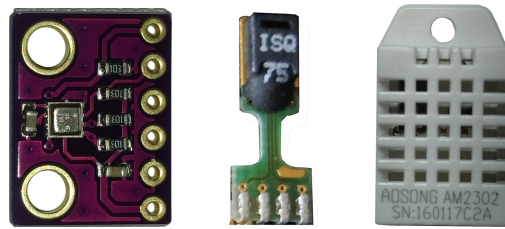


ABBILDUNG 3.5: Temperatur- / Luftfeuchtigkeitssensoren (von links): BME280, SHT75, DHT22.

Feinstaub

Konzentrationen von Feinstaubpartikeln können mit Sensoren gemessen werden. Besonders einfache Partikelsensoren, für größere Rauchpartikel, sind in Rauchmeldern enthalten. In dieser Arbeit wurde der kostengünstige Feinstaubsensor SDS011 in den Gas-Sensorboxen in Abschnitt 3.5.5 eingesetzt. Bei dieser Sensorbauweise wird die Außenluft mit einem Lüfter in eine Kammer gesaugt und das Streulicht eines Lasers gemessen (Laser-Scattering) (Budde et al., 2018). Die bei Feinstaub am häufigsten in Literatur und anderen Quellen angegebenen Partikelgrößen sind $2,5\text{ }\mu\text{m}$ und $10\text{ }\mu\text{m}$. Bei vielen Sensoren sind allerdings aufgrund ihrer Messmethode bei $10\text{ }\mu\text{m}$ auch die Konzentrationen kleinerer Partikel beinhaltet. Einige Einschränkungen des SDS011-Sensors sind in der Veröffentlichung von Budde et al. (2018) aufgelistet. Hierzu zählen beispielsweise die Feuchtigkeitsabhängigkeit und Abweichungen zwischen Sensoren desselben Typs.

Kontaktsensorik

Eine sehr einfache Art von Sensor sind Kontaktsensoren, die binär zwei Zustände unterscheiden können. Im Smarthome werden diese häufig an Fenstern und Türen angebracht, um deren Zustände „offen“ und „geschlossen“ erfassen zu können. Typischerweise werden diese Sensoren in Smarthomes bei Sicherheitssystemen verwendet. Sie können jedoch auch dafür genutzt werden, die Heizung abzuschalten, wenn ein Fenster zum Lüften geöffnet ist, um Energie zu sparen. Generell existieren jedoch eine Vielzahl weiterer Einsatzzwecke.



ABBILDUNG 3.6: Funk-Kontaktsensor des Homematic-Systems. Der optische Sensor befindet sich unten.

Für die Erkennung der Zustände können verschiedene Methoden eingesetzt werden. Im Folgenden werden drei Möglichkeiten vorgestellt. Eine einfache Erfassungsmethode stellt ein mechanischer Mikroschalter dar, bei dem durch Druck ein Kontakt geschlossen wird. Eine weitere mechanische Methode ist die Nutzung von Reedschaltern, die durch einen Magneten geschaltet werden. Eine nichtmechanische Methode ist mit der Nutzung optischer Sensorik möglich. Hierbei wird ein Sender und Empfänger von infrarotem Licht genutzt. In Abbildung 3.6 ist ein optischer Funk-Kontaktsensor des Homematic-Systems abgebildet. Das Licht wird dabei vom zu erkennenden Objekt (z. B. eine Tür-Kante) reflektiert, wenn dieses sich vor dem Sensor befindet. Bei dunklen Farben, wie z. B. schwarz, welches das Licht weitestgehend absorbiert, muss eine reflektierende Folie (z. B. weiß oder Aluminium) aufgeklebt werden, um einen geschlossenen Zustand erkennen

zu können. Ein vergleichbares Prinzip wird bei einer Reflexions-Lichtschranke, bei deutlich höherer Reichweite und relativ hohem Energieaufwand, genutzt.

Neigungssensorik

Mit Neigungssensoren kann beispielsweise erkannt werden, ob ein Kipptor einer Garage offen oder geschlossen ist. Bei einfachen mechanischen Sensoren befindet sich eine kleine, frei bewegliche Metallkugel in einer Röhre. Ab einem bestimmten Neigungswinkel rollt die Kugel auf die eine oder andere Seite der Röhre, wodurch ein Kontakt geschlossen wird. Früher wurde hierfür das bei Raumtemperatur flüssige Metall „Quecksilber“ eingesetzt, das jedoch stark toxisch ist. Da die Erdgravitation (Erdbeschleunigung) genutzt wird, um die Neigung festzustellen, können alternativ die präziseren Beschleunigungssensoren verwendet werden, die jedoch teurer sind.

Energieverbrauchssensorik

Mit dem gemessenen Strom I in Ampere und der Spannung U in Volt kann die elektrische Leistung P in Watt mit der Formel $P = UI$ ermittelt werden (Hering et al., 2016, S. 258). Die elektrische Arbeit W in Wattsekunden, die in einer bestimmten Zeitspanne t in Sekunden verrichtet wurde, wird folgendermaßen berechnet: $W = Pt$ (Hering et al., 2016, S. 258).

Ein elektrisches Gerät kann beispielsweise elektrische Energie in mechanische Arbeit umwandeln, indem es elektrische Arbeit verrichtet. Für den physikalischen Begriff „elektrische Arbeit“ werden in einer großen Anzahl an Fachliteratur und in der Umgangssprache die Begriffe „Stromverbrauch“ oder „Energieverbrauch“ synonym verwendet.

Um den Stromverbrauch elektrischer Geräte zu überwachen, ist es beispielsweise möglich Funk-Zwischenstecker zu nutzen. Diese werden häufig Energiemessgeräte oder Strommessgeräte genannt. Viele zu überwachende Geräte zeigen, je nach Typ und Nutzung, charakteristische Muster in ihrem Stromverbrauch, die unterschieden werden können. Im einfachsten Fall können anhand des aktuellen Verbrauchs die Zustände „an“ und „aus“ erkannt werden.

Schallsensorik

Gemäß Hering et al. (2016, S. 553) wird Schall je nach Frequenzverlauf des Schalldrucks in Ton, Geräusch oder Knall unterteilt. Gemäß den Autoren besitzt ein Ton nur eine Schallfrequenz und einen sinusförmigen Druckverlauf. Ein Geräusch verfügt hingegen über ein breitbandiges Frequenzspektrum und einen stark schwankenden Druckverlauf. Die Autoren definieren Schall als Knall, wenn bei einem sehr breitbandigen Frequenzspektrum ein nahezu konstanter Druckverlauf auftritt.

Mikrofone sind Schallwandler, die Schalldruck in elektrische Spannung umwandeln. Hierbei existieren verschiedene Wandlerprinzipien. Beispielsweise basieren Kondensatormikrofone und Elektretmikrofone auf dem elektrostatischen Prinzip, bei dem sich die elektrische Kapazität eines Kondensators, bestehend aus einer Schallwandlarmembran und Gegenelektrode, ändert (Hering et al., 2016, S. 559). Mikrofone, die auf dem elektrostatischen Prinzip basieren, besitzen eine hohe Empfindlichkeit, flachen Frequenzgang und geringen Rauschpegel (Shah et al., 2019). Moderne Mikrofone können als mikroelektromechanische Systeme (MEMS) gefertigt werden. Sogenannte „MEMS-Mikrofone“ besitzen im Vergleich zu den bereits kompakten Elektretmikrofonen eine wesentlich geringere Bauteilgröße von nur wenigen Quadratmillimetern und reagieren weniger empfindlich auf Temperaturänderungen sowie Vibrationen (Shah et al., 2019). Da MEMS-Mikrofone zudem einen geringen Preis besitzen, werden sie in einer Vielzahl von IoT- und Mobile-Computing-Geräten wie Smartspeakern und Smartphones integriert.

Eine einfache Auswertung der relativen Lautstärke ermöglicht die Erkennung von Schallquellen in der Umgebung, die lauter als das Rauschen sind. Komplexere Auswertungen, wie die Identifikation von bestimmten Geräuschquellen, sind nur im Frequenzbereich möglich. Im mobilen Einsatz

muss beachtet werden, dass viele unerwünschte Geräusche z. B. durch Bewegungsgeräusche, entstehen. Diese müssen gefiltert oder durch eine geeignetere Mikrofon-Position reduziert werden.

Masse- und Volumenssensorik

Die Masse von Objekten kann mittels verschiedener Sensoren bestimmt werden. Die Masse m in Kilogramm wird in der Regel über die gemessene Gewichtskraft F in Newton und die Erdbeschleunigung ($g = 9,81 \text{ m s}^{-2}$) mit $m = F/g$ ermittelt (Hering et al., 2016, S. 35). Eine Berechnung des Volumens V in Kubikmetern ist über die Dichte ρ in kg m^{-3} und die berechnete Masse m in Kilogramm möglich: $V = m/\rho$ (Hering et al., 2016, S. 110). Getränke bestehen meist zu einem Großteil aus Wasser. Daher kann deren Masse vereinfacht mit der Dichte von Wasser $\rho_{\text{Wasser}} = 998 \text{ kg m}^{-3}$ (Hering et al., 2016, S. 221) in ihr Volumen umgerechnet werden. Ein Kilogramm Wasser entspricht daher ca. einem Liter.

Zur Ermittlung der Masse können kalibrierte Drucksensoren oder Wägezellen (engl. load cell) genutzt werden. Im Vergleich zu Drucksensoren liefern Wägezellen präzisere Werte und besitzen über die Zeit einen geringeren Sensordrift. Stärkere Änderungen der Umgebungstemperatur müssten allerdings für eine exakte Messung kompensiert werden. Das Diagramm in Abbildung 3.8 zeigt das über mehrere Stunden mit einer Wägezelle gewogene Gewicht einer mit unterschiedlichen Flüssigkeitsmengen gefüllten PET-Flasche. Die Messwerte wurden nicht gefiltert, dennoch sind keine Störungen ersichtlich.

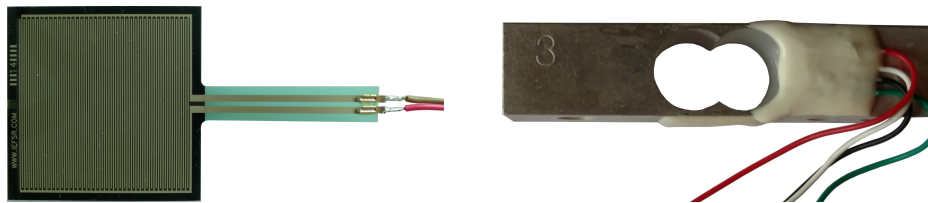


ABBILDUNG 3.7: Drucksensor (links) und Wägezelle (rechts).

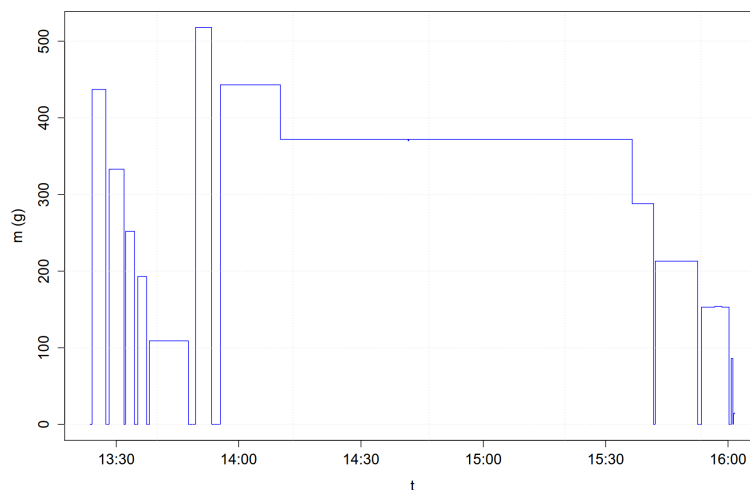


ABBILDUNG 3.8: Sensordaten einer Wägezelle (maximale Traglast: 5 kg) mit verschiedenen Flüssigkeitsmengen in einer PET-Flasche.

Gewichtsverteilungssensorik

Druckänderungen über Flächen können mit drucksensitiven Matten gemessen werden, die entweder resistiv oder kapazitiv arbeiten. Bei resistiven Sensoren ändert sich der gemessene Widerstand

abhängig vom Druck. Bei kapazitiven Sensoren variiert die Kapazität, die mittels eines einfachen Schwingkreises bestimmt wird. Die Veränderung der Kapazität wird auch bei modernen Touchscreens genutzt, um die Position der Finger auf dem Display zu bestimmen.

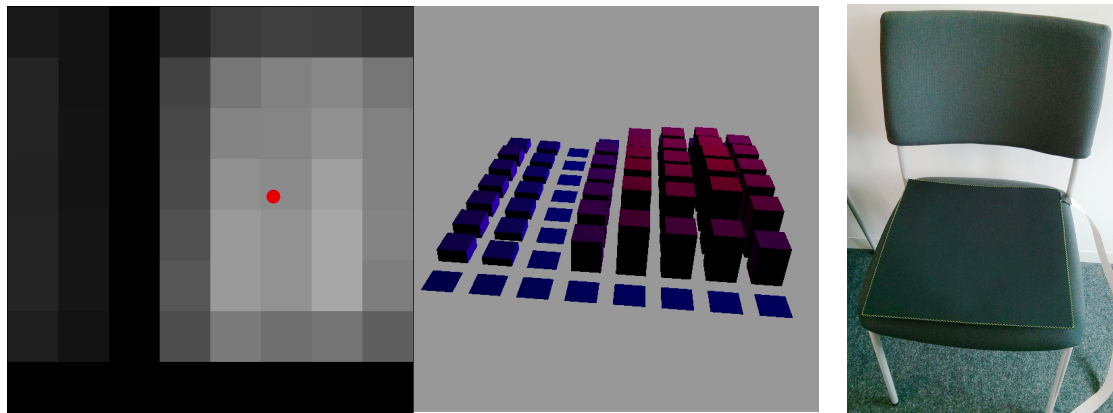


ABBILDUNG 3.9: Druckmatrix einer resistiven Druckmatte (SensingTex) mit einer Auflösung von 8x8. Der Druckmittelpunkt (engl. Center of Pressure) ist hier mit einem roten Punkt markiert. Die zugehörige Druckmatte auf einem Stuhl (ohne Person) ist ebenfalls abgebildet.

In Abbildung 3.9 ist die Druckverteilung beim Sitzen auf einer resistiven Druckmatte und die zugehörige Matte, auf einem Stuhl liegend, dargestellt. Sehr einfache Sensoren, wie eine Tanzmatte, liefern ausschließlich einen binären Wert zurück, ob auf einer abgegrenzten Fläche ein bestimmter Druck vorhanden ist. Je nach Auflösung der Matte ist es z. B. möglich kleinere Objekte oder beim Menschen einzelne Füße zu unterscheiden, auch wenn sich diese nicht auf festgelegten Positionen befinden.

Abstandssensorik

Um zu erkennen, wie weit eine Person oder ein Objekt von einem definierten Punkt entfernt ist, können optische oder akustische Abstandssensoren eingesetzt werden. Diese werden beispielsweise für berührungslose Seifenspender (Infrarot) oder für Einparkhilfen bei Autos (Ultraschall) eingesetzt. Abstandssensoren senden ein optisches (Infrarot) bzw. akustisches (Ultraschall) Signal aus, das vom zu messenden Objekt reflektiert und von einem Empfänger erfasst wird. Zur Bestimmung des Abstands wird die Zeitspanne zwischen dem Senden und Empfangen gemessen (Time of Flight = ToF). Die Lichtgeschwindigkeit (im Vakuum: ca. $299\,792\,\text{km s}^{-1}$ (Hering et al., 2016, S. 776)) bzw. Schallgeschwindigkeit (in Luft: ca. $344\,\text{m s}^{-1}$ (Hering et al., 2016, S. 556)) sind hierbei bekannt. Einfachere Infrarot-Entfernungssensoren messen nicht die bei Licht sehr kurzen Zeitspannen, sondern die Intensität des ausgesendeten, reflektierten Lichts. Die Sensoren verfügen über Vor- und Nachteile. Bei optischen Sensoren kann beispielsweise die Farbe „Schwarz“ zu falschen Abstandsmessungen führen, wenn das ausgesandte Licht stark absorbiert wird. Zudem können Reflexionen an Glasscheiben Störungen verursachen. Bei akustischen Sensoren treten vergleichbare Probleme auf, falls der ausgesandte akustische Impuls von einer Oberfläche stark absorbiert wird oder zusätzliche Schallreflexionen auftreten.

In dieser Arbeit wurden zwei optische Abstandssensoren eingesetzt: Sharp GP2Y0A02YK0F¹⁴ (Erkennungsbereich: 20 - 150 cm) und ST VL53L0X¹⁵ (Erkennungsbereich bis 200 cm). Der Letztere ist hierbei ein aktuelleres Model, das einen IR-Laser und ToF einsetzt und zudem wesentlich

¹⁴https://www.sharpsde.com/fileadmin/products/Optoelectronics/Sensors/Specs/GP2Y0A02YK0F_13May05_Spec_ED-05G127.pdf (21.02.2021)

¹⁵<http://www.st.com/en/imaging-and-photonics-solutions/vl53l0x.html> (10.02.2020)

kompakter ist. Zusätzlich wurde ein Ultraschall-Abstandssensor vom Typ HC-SR04¹⁶ (Erkennungsbereich: 2 - 400 cm) genutzt. Die Sensoren sind in Abbildung 3.10 dargestellt.

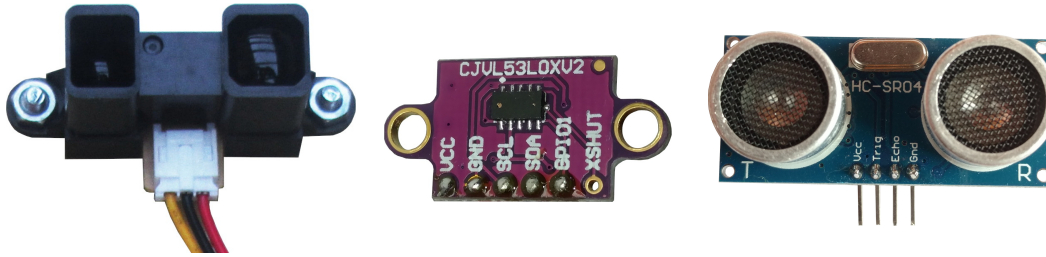


ABBILDUNG 3.10: Von links: Entfernungssensoren Sharp GP2Y0A02YK0F (optisch), VL53L0X (optisch), HC-SR04 (akustisch).

3.5 Überblick Gassensorik

Mithilfe von Gassensoren ist es möglich, mittels verschiedenen Messprinzipien die Konzentration von einzelnen Gasen oder Gasgemischen zu erfassen. Die Konzentrationen verschiedener Gase sind charakteristisch für verschiedene Stoffe, von denen sie abgegeben werden. In diesem Kapitel werden zunächst verschiedene Messprinzipien vorgestellt, wobei insbesondere Prinzipien, die in eigenen Arbeiten genutzt wurden, ausführlicher behandelt werden. Es handelt sich bei diesen um kostengünstigere Sensorik, die in näherer Zukunft auch im Smarthome-Bereich oder bei mobilen Geräten weitere Verbreitung finden kann.

3.5.1 Messprinzipien

Überblick

Über aktuelle Messprinzipien gibt die Arbeit von Nazemi et al. (2019) einen Überblick. Die folgenden Prinzipien werden in dem Artikel aufgeführt, wobei das NDIR-Prinzip zusätzlich ergänzt wurde.

- CNT (Carbon Nano-Tubes)
- Akustisch
 - QCM (Quartz Crystal Micro-balance)
 - SAW (Surface Acoustic Waves)
 - CMUT (Capacitive Micro-machined Ultrasonic Transducer)
- Elektrochemisch
 - MOS (Metal Oxide Semiconductor)
- Optisch
 - Fiber-Optic
 - PhC (Photonic Crystal)
 - NDIR (Non-Dispersive InfraRed)

Im Folgenden werden die Prinzipien „MOS“ (Metal Oxide Semiconductor) und „NDIR“ (nicht-dispersive IR-Spektroskopie oder Non-Dispersive InfraRed) genauer betrachtet.

¹⁶<https://cdn.sparkfun.com/datasheets/Sensors/Proximity/HCSR04.pdf> (21.02.2021)

MOS-Prinzip

Eine Beschreibung des elektrochemischen MOS-Prinzips ist im Artikel von Nazemi et al. (2019) enthalten. Beim MOS-Prinzip findet eine Änderung der Leitfähigkeit einer Metalloxidschicht statt, sobald ein Zielgas mit dieser reagiert. Entsprechend dem zu detektierenden Zielgas wird das Material der Schicht gewählt. Damit eine Reaktion zwischen der Metalloxidschicht und dem Zielgas stattfinden kann, muss die Schicht mit einem Heizelement auf Temperaturen zwischen 150 und 400 °C erhitzt werden Nazemi et al. (2019).

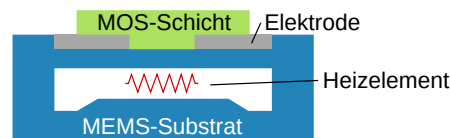


ABBILDUNG 3.11: Schematischer Querschnitt durch einen MOS-Gassensor, der in kompakter MEMS-Fabrikationstechnik produziert wurde. Abbildung nach Nazemi et al. (2019).

MOS-Sensoren können als mikroelektromechanisches System hergestellt werden. Hierdurch kann das MOS-Prinzip mit weniger Stromverbrauch, wesentlich kompakter genutzt werden. Ein Einsatz in mobilen Geräten oder kompakten, batteriebetriebener Sensorik im Smarthome wird in Form eines MEMS-Sensors möglich.

NDIR-Prinzip

Das Prinzip der optischen nichtdispersiven IR-Spektroskopie und der Aufbau eines Sensors wird nachfolgend gemäß der Arbeit von Mendes et al. (2015) beschrieben. Eine schematische Darstellung eines NDIR-Sensors ist in Abbildung 3.12 dargestellt. In diesem Fall handelt es sich um einen einfachen Sensor, der nur einen IR-Sensor besitzt. NDIR-Sensoren bestehen aus einer Messröhre mit einer Infrarot-Lichtquelle, einem optischen Filter und einem Lichtsensor. Es wird ein Filter genutzt, der nur eine spezifische Wellenlänge der IR-Lichtquelle durchlässt, die vom zu bestimmenden Gas absorbiert wird. Bei Kohlendioxid werden beispielsweise die Wellenlängen 4,2 µm und 14,9 µm absorbiert, wobei CO₂ gut für dieses Messprinzip geeignet ist, da nur wenige Interferenzen mit anderen Gasen auftreten (Hanst, 1986). Ein geeigneter Filter ist somit beim Zielgas CO₂ ausschließlich für die Wellenlänge 4,2 µm durchlässig. Hinter dem Filter befindet sich ein Sensor, der die Lichtintensität misst. Die gemessene Intensität ist proportional zur Konzentration des Zielgases.

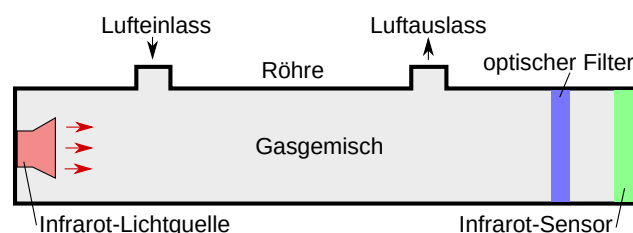


ABBILDUNG 3.12: Schematischer Querschnitt durch einen einfachen Gassensor nach dem NDIR-Prinzip. Abbildung nach Mendes et al. (2015).

3.5.2 Sensorpositionierung

Die Positionierung von Gassensorik hängt von den zu messenden Gasen und den räumlichen Verhältnissen ab. Gase steigen, je nach dem Gewicht der Moleküle, in der Luft auf, sinken oder schweben. Zudem können sie durch Luftströmungen aufgrund von Thermik oder Belüftung im Raum ungleichmäßig verteilt sein.

Eine Übersicht über die Dichte verschiedener Gase ist in Tabelle 3.2 dargestellt. Bei Normbedingungen besitzen 1 mol eines idealen Gases ein Volumen von $V_m = 22,41 \text{ mol}^{-1}$ (Meschede, 2015, S. 104). Die relative Atommasse (in der Einheit u) von Gasen kann direkt im Periodensystem abgelesen werden und entspricht der molaren Masse M in g mol^{-1} . Somit kann das Gewicht eines Gasmoleküls pro Liter (Dichte) berechnet werden, indem für die Moleküle die molaren Massen der Atome zunächst addiert und daraufhin durch 22,4 geteilt werden. Beispielsweise besitzt ein CO_2 -Molekül die molare Masse: $M_{\text{CO}_2} = (12,0 + 2 \cdot 16,0) \text{ g mol}^{-1} = 44,0 \text{ g mol}^{-1}$. Ein Liter des CO_2 -Gases besitzt somit eine Masse von $m_{\text{CO}_2} = 1 \text{ l} \cdot 44,0 \text{ g mol}^{-1} / 22,41 \text{ mol}^{-1} \approx 1,96 \text{ g}$. Mit der Dichte von trockener Luft ($\rho = 28,97 \text{ g mol}^{-1}$ ¹⁷) kann abgeschätzt werden, wie sich ein Gas in Luft verhält und wo der entsprechende Sensor platziert werden sollte oder ob eine aktive Belüftung des Sensors notwendig ist.

Summenformel	Gas	g/l	Summenformel	Gas	g/l
H_2	Wasserstoff	0,09	NO	Stickstoffmonoxid	1,34
CH_4	Methan	0,71	C_3H_8	Propan	1,96
NH_3	Ammoniak	0,76	CO_2	Kohlendioxid	1,96
CO	Kohlenmonoxid	1,25	$\text{C}_2\text{H}_5\text{OH}$	Alkohol	2,05
	Luft (trocken)	1,29	NO_2	Stickstoffdioxid	2,05
			C_4H_{10}	Iso-Butan	2,59

TABELLE 3.2: Übersicht über die Dichte verschiedener Gase. Gase mit geringerer Dichte als Luft steigen auf (links), mit höherer Dichte sinken sie ab (rechts). Die Dichte der zu messenden Gase spielt daher eine bedeutsame Rolle bei der Positionierung von Sensoren.

Je nach Größe des Raums, zu erwartenden Luftströmungen und der Art der zu messenden Gase kann es nötig sein an verschiedenen Positionen und Höhen Gassensoren zu platzieren. Für Kohlendioxid ist es beispielsweise sinnvoll, den Sensor nicht am Boden (höchste Konzentration) oder der Decke (geringste Konzentration) zu positionieren. Zur Abschätzung der CO_2 -Exposition von Erwachsenen und Kindern in einem Raum, die sitzen oder stehen, ist beispielsweise eine Höhe von ca. einem Meter über dem Boden geeignet.

3.5.3 Herausforderungen bei der Auswertung

Bei der Auswertung von Gassensordaten können verschiedene Probleme auftreten, die berücksichtigt oder kompensiert werden müssen.

Verzögerungen

Gassensoren reagieren verzögert auf Änderungen von Gaskonzentrationen. Dies liegt einerseits daran, dass diese von der Position im Raum und möglichen Luftströmungen abhängen. Andererseits besitzen die verschiedenen eingesetzten Messprinzipien und technischen Umsetzungen in Sensoren unterschiedliche Ansprechzeiten und Verzögerungen. Beispielsweise benötigt bei NDIR-Sensoren das Einströmen des Gasgemisches in die Messröhre eine gewisse Zeit (Zakaria, 2010, S. 43).

Sensordrift

Wie auch bei einigen anderen Arten von Sensoren, treten bei den Messwerten von Gassensoren über längere Zeitspannen hinweg Abweichungen der Grundlinie (engl. baseline drifts) auf (Piedrahita et al., 2014). Laut Piedrahita et al. (2014) entstehen diese beim NDIR-Prinzip weniger ausgeprägt als beim MOS-Prinzip. Die Autoren identifizierten zwei wesentliche Gründe, weshalb Abweichungen

¹⁷<https://nssdc.gsfc.nasa.gov/planetary/factsheet/earthfact.html> (05.06.2019)

bei MOS-Sensoren auftreten können: 1) die Wärmeabgabe des Heizelements verändert sich, 2) eine „Vergiftung“ oder dauerhafte chemische Verbindung mit der Metalloxidschicht hat stattgefunden. Eine „Vergiftung“ kann im größeren Umfang zur Zerstörung des Sensors führen, da anschließend eine Reaktion von Gasen mit der Sensorschicht nur noch begrenzt oder überhaupt nicht mehr möglich ist. Auch bei NDIR-Sensoren kann eine Abweichung über die Zeit hinweg auftreten, falls beispielsweise die Helligkeit der verwendeten Lichtquelle abnimmt.

Abhängigkeit von anderen Gasen

Insbesondere MOS-Sensoren reagieren nicht ausschließlich auf ein einziges Zielgas. Daher kann nicht ausgeschlossen werden, dass andere Gase Sensorwerte beeinflussen. Die Reaktion mit anderen Gasen hängt vom Material der verwendeten Metalloxidschicht ab. Bei Zinnoxid (SnO_2) ist beispielsweise eine Reaktion mit Methan (CH_4) und Kohlenmonoxid (CO) bekannt (Ionescu et al., 1999). Weitere Abhängigkeiten können in der Regel in den Datenblättern der entsprechenden Sensoren gefunden werden. Beim NDIR-Prinzip tritt je nach zu messendem Gas eine Abhängigkeit auf, falls mehrere Gase die gleichen Wellenlängen absorbieren. Wie bereits in Abschnitt 3.5.1 erläutert, ist bei Kohlendioxid allerdings kaum eine Interferenz mit anderen Gasen zu erwarten.

Abhängigkeit von Temperatur und Luftfeuchtigkeit

MOS-Sensoren besitzen Abhängigkeiten hinsichtlich Temperatur und der Luftfeuchtigkeit. In den Arbeiten von Ionescu et al. (1999) und Barsan und Weimar (2003) wurde ausführlich die starke Abhängigkeit der Reaktion des Metalloxids SnO_2 von der Luftfeuchtigkeit untersucht. Die Reaktion von Gasen mit der Metalloxidschicht hängt zusätzlich von der Temperatur ab. In der Regel wird daher eine hohe Betriebstemperatur zwischen 200 und 400 °C (Barsan & Weimar, 2003) genutzt, um diesen Einfluss zu verringern. Auch in den Datenblättern kostengünstiger MOS-Sensoren (z. B. der MQ-Reihe) ist eine deutliche Abhängigkeit der Messwerte bezüglich der Temperatur und Luftfeuchtigkeit ersichtlich. NDIR-Sensoren können, je nach eingesetzter Lichtwellenlänge, ebenfalls von der Luftfeuchtigkeit beeinflusst werden (Zakaria, 2010, S. 43). Einige Sensorgeräte führen bereits per Firmware eine automatische Kompensation durch.

Aufwärmphase

Während der Aufwärmphase liefern viele Gassensoren keinen stabilen Sensorwert. Bei MOS-Sensoren muss zunächst die Zieltemperatur des Heizelements erreicht werden. Daher sollten erste Messwerte erst nach einer entsprechenden Zeitdauer ermittelt werden. Bei niedrigen Umgebungstemperaturen kann die benötigte Aufwärmphase länger ausfallen. NDIR-Sensoren besitzen messprinzipbedingt keine Aufwärmphase.

Einzigartiges Sensorverhalten

MOS-Sensoren weisen einzigartige Unterschiede untereinander auf, die durch die Fertigung entstehen. Laut L. Zhang et al. (2014) zählen hierzu der Grundwert- und Empfindlichkeits-Unterschied. Abbildung 3.13 zeigt die synchronisierten Sensordaten von zwei MQ2-Sensoren, die denselben Gaskonzentrationen ausgesetzt waren. Die Daten wurden mit den zwei stationären Gas-Sensorboxen aufgezeichnet (siehe Abschnitt 3.5.5). Es ist deutlich der Unterschied zwischen den Ausgaben der beiden Sensoren erkennbar. Primär sind hier ein stabiler Offset (Baseline) der Spannungsverläufe und ein unterschiedlicher Wertebereich (Empfindlichkeit) zu erkennen.

3.5.4 Verwendete Sensoren

In diesem Abschnitt werden Gassensoren vorgestellt, die in dieser Arbeit verwendet wurden.

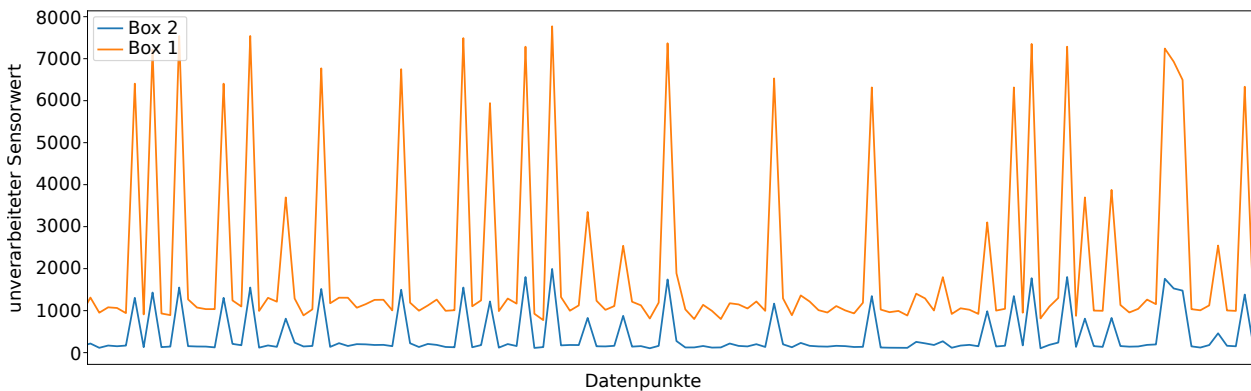


ABBILDUNG 3.13: Synchronisierte Daten von zwei MQ2-Gassensoren, die sich im selben Verhältnis befanden. Es zeigen sich für MOS-Sensoren typische Unterschiede beim Grundwert und bei der Empfindlichkeit.

MOS-Prinzip

MQ-Reihe In dieser Arbeit wurden mehrere Sensoren der MQ-Reihe in Untersuchungen eingesetzt. In Abbildung 3.14 sind zwei MQ-Sensoren mit Ansteuerungsplatinen dargestellt. Die Schaltungen ermöglichen es, die Stromversorgung mit zwei Anschlüssen herzustellen und die sich abhängig von Gaskonzentrationen verändernde Ausgangsspannung der Sensoren abzugreifen. Diese Spannung kann anschließend mit einem Analog-Digital-Wandler digitalisiert werden.

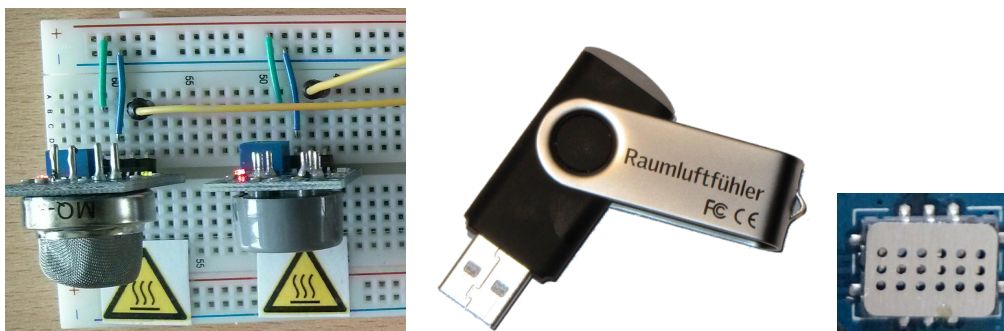


ABBILDUNG 3.14: Links: Zwei verschiedene Typen von MQ-Sensoren auf einem Steckbrett. Mitte: Variante des iAM USB-Sticks zur Messung von VOCs. Rechts: Verlöteter MiCS-6814 Gassensor.

Sensoren der MQ-Reihe sind sehr preisgünstig und können bereits für ca. 1 € auf Handelsplattformen erworben werden. Sie besitzen allerdings einen hohen Stromverbrauch. Gemäß Datenblatt benötigt beispielsweise der MQ3-Sensor eine elektrische Leistung von ca. 750 mW (siehe Tabelle 3.3). Wegen des hohen Stromverbrauchs und der Größe ist eine mobile oder batteriebetriebene Nutzung nur eingeschränkt möglich. Bei einer Akkukapazität von 12 Wh (in dieser Größenordnung liegen viele Smartphone-Akkus) wäre beispielsweise bei Dauerbetrieb eine theoretische Laufzeit von 16 Stunden zu erwarten, ohne dass zusätzlich benötigte Komponenten, wie ein Mikrocontroller oder Funkmodule (z. B. Bluetooth oder WiFi) einberechnet wurden. Eine temporäre Abschaltung des Sensors bedeutet eine erneute Aufwärmphase, sodass dies nur begrenzt zur Energieeinsparung genutzt werden kann.

Die verschiedenen Modelle der MQ-Reihe reagieren auf unterschiedliche Zielgase und unterscheiden sich auch hinsichtlich ihres Stromverbrauchs. Als Metalloxid kommt bei allen angegebenen MQ-Sensoren Zinkoxid (SnO_2) zum Einsatz. In Tabelle 3.3 ist eine Übersicht dargestellt, deren

Daten aus folgenden Datenblättern entstammen: MQ2¹⁸, MQ3¹⁹, MQ5²⁰, MQ9²¹.

Name	Gase	Metallox.	Leistung
MQ2	LPG, Iso-Butan, Propan, Methan, Alkohol, Wasserstoff	SnO ₂	800 mW
MQ3	Alkohol, Benzin	SnO ₂	750 mW
MQ5	Wasserstoff, LPG, Methan, Kohlenmonoxid, Alkohol	SnO ₂	800 mW
MQ9	LPG, Kohlendioxid, Methan	SnO ₂	340 mW

TABELLE 3.3: Übersicht über in dieser Arbeit genutzte MQ-Sensoren. Die Daten entstammen den im Text angegebenen Datenblättern.

iAM Als Sensorgerät für flüchtige organische Verbindungen (engl. volatile organic compounds: VOCs) wurde der AMS iAM USB-Stick²² verwendet. Dieser wird unter verschiedenen Bezeichnungen von unterschiedlichen Firmen vertrieben. Der in den Stick integrierte Gassensor basiert auf dem MOS-Prinzip, das für die Erkennung von VOCs bevorzugt eingesetzt wird (Nazemi et al., 2019). Der Website kann entnommen werden, dass er insgesamt eine Leistung von 150 mW benötigt. Das gesamte Sensorgerät kostet ca. 25 €. Abbildung 3.14 zeigt eine Variante des USB-Sticks.

MiCS-6814 Der MiCS-6814²³ (siehe Abbildung 3.14) Gassensor enthält drei verschiedene Sensorelemente, die nach dem MOS-Prinzip arbeiten. Der Sensor ist sehr kompakt (5×7×1,55 mm), da er als modernes mikroelektromechanisches System gefertigt wurde (siehe Datenblatt). Laut Datenblatt erfasst der „OX“ Chip oxidierende Gase, der „RED“ Chip reduzierende Gase und der „NH3“ Chip primär Ammoniak. Der gesamte Sensor benötigt eine Leistung von ca. 100 mW.

In dieser Arbeit wurde der MiCS-6814 Sensor in Form einer Fertigplatine mit Mikrocontroller (10 bit AD-Wandler) und Firmware genutzt: ²⁴. Der Preis der kompletten Platine beträgt ca. 40 €. Die Firmware liefert neben den Rohwerten auch vom Hersteller kalibrierte Werte für bestimmte Gase, die aus den Rohwerten der drei Sensorelemente errechnet wurden.

Die Sensorelemente-Kombination kann laut Datenblatt zur Erfassung folgender Gase genutzt werden: Kohlenmonoxid, Stickstoffdioxid, Alkohol, Wasserstoff, Ammoniak, Methan, Propan und Iso-Butan. Im Vergleich zur MQ-Reihe ist dieser modernere Sensor wesentlich kompakter und besitzt einen deutlich geringeren Stromverbrauch, wodurch er auch für mobile Geräte geeignet ist.

NDIR-Prinzip

NDIR-Sensoren wurden in dieser Arbeit ausschließlich zur Messung der Kohlendioxidkonzentration eingesetzt. Die Ausgabedaten beider Sensorgeräte sind für eine direkte Auswertung geeignet.

AirCO2ntrol Mini Beim TFA AirCO2ntrol-Mini-Sensorgerät²⁵ handelt es sich um einen Kohlendioxidssensor basierend auf dem NDIR-Prinzip. Die fertig aufgebaute Schaltung mit Gehäuse beinhaltet neben dem CO₂-Sensor einen Temperatur-Sensor und zur Anzeige der Messwerte ein LCD-Display sowie Indikator-LEDs zur Bewertung der aktuellen CO₂-Konzentration (Luftqualitätsampel). Das Gerät kann die Sensordaten zusätzlich per USB an einen Rechner übermitteln. Gemäß der Hersteller-Website liegt der Messbereich der CO₂-Konzentration zwischen 0 und 3000 ppm. Die benötigte elektrische Leistung liegt laut Bedienungsanleitung bei durchschnittlich 100 mW.

¹⁸https://seeddoc.github.io/Grove-Gas_Sensor-MQ2/res/MQ-2.pdf (14.10.2020)

¹⁹https://seeddoc.github.io/Grove-Gas_Sensor-MQ3/res/MQ-3.pdf (14.10.2020)

²⁰https://seeddoc.github.io/Grove-Gas_Sensor-MQ5/res/MQ-5.pdf (14.10.2020)

²¹https://seeddoc.github.io/Grove-Gas_Sensor-MQ9/res/MQ-9.pdf (14.10.2020)

²²<https://ams.com/iam> (27.06.2019)

²³https://www.sgxsensortech.com/content/uploads/2015/02/1143_Datasheet-MiCS-6814-rev-8.pdf (28.06.2019)

²⁴http://wiki.seeedstudio.com/Grove-Multichannel_Gas_Sensor/ (28.06.2019)

²⁵<https://www.tfa-dostmann.de/produkt/co2-monitor-airco2ntrol-mini-31-5006/> (28.06.2019)

Die Front des Sensorgeräts ist in Abbildung 3.15 dargestellt. Das Gerät wird von verschiedenen Resellern wie TFA angeboten.

SCD30 Der Sensirion SCD30-Kohlendioxidsensor²⁶ kann in eigene Schaltungen integriert werden. Abbildung 3.15 zeigt die SMD bestückte Seite des Sensors. Die Daten können per I²C, UART oder PWM ausgelesen werden. Er nutzt ebenfalls das NDIR-Prinzip und besitzt zusätzlich einen Sensor zur Messung der Temperatur und Luftfeuchtigkeit. Hierbei ist der Sensor vergleichsweise kompakt. Eine Firmware führt eine automatische Kompensation gegenüber variierender Temperatur und Luftfeuchtigkeit durch. Optional ist dies auch hinsichtlich des Luftdrucks möglich, wofür jedoch ein eigener Sensor benötigt wird. Der Messbereich der CO₂-Konzentration liegt laut Datenblatt zwischen 400 und 10.000 ppm. Es ist eine benötigte mittlere elektrische Leistung von ca. 63 mW (19 mA bei 3,3 V) bei einem Messintervall von 2 Sekunden angegeben.



ABBILDUNG 3.15: Links: AirCO2ntrol Mini Kohlendioxidsensorgerät. An der rechten Seite ist das USB-Kabel eingesteckt. Rechts: SCD30-Kohlendioxidsensor mit angelöteter Stiftleiste.

3.5.5 Prototypen

Stationäre Sensorboxen

Zur Messung verschiedener Gase wurden für mehrere Untersuchungen zwei Sensorboxen-Prototypen mit identischer Sensorik entwickelt. Die Sensoren befanden sich hierbei in je einem 3D-gedruckten Gehäuse, der jeweilige ansteuernde / auslesende Raspberry Pi 3 B außerhalb davon. Eine Sensorbox ist in Abbildung 3.16 ohne Deckel dargestellt. Tabelle 3.4 gibt einen Überblick über die enthaltenen Komponenten. Die beiden per USB angeschlossenen Sensoren (AirCO2ntrol Mini, iAM) waren optional. Der Lichtsensor wurde hinzugefügt, um beispielsweise erkennen zu können, ob Sonnenlicht die Sensorbox erwärmt. Insgesamt benötigte eine Sensorbox eine Leistung von ca. 3,3 W. Der daran angeschlossene Raspberry eine Leistung von ca. 1,5 W.

Mobile Sensorbox

Eine mobile, akkubetriebene Gas-Sensorbox (siehe Abbildung 3.16) wurde entwickelt, um Gassen-sorik ohne feste Verkabelung in Experimenten nutzen zu können. In diesem Fall mussten die Größe und der Stromverbrauch besonders berücksichtigt werden. In Tabelle 3.4 sind die Komponenten aufgelistet. Zur Stromversorgung war in dem 3D-gedruckten Gehäuse eine USB-Powerbank mit 5,2 Ah verbaut. Mit aktiven Sensoren und WiFi-Verbindung wurden bei einer Spannung von 5 V ca. 280 mA Strom verbraucht (Leistung ca. 1,4 W), wodurch eine geschätzte, maximale Laufzeit von ca. 18 h erreicht werden konnte.

²⁶https://www.sensirion.com/fileadmin/user_upload/customers/sensirion/Dokumente/9.5_CO2/Sensirion_CO2_Sensors_SCD30_Datasheet.pdf (28.06.2019)

stationäre Gas-Sensorbox		mobile Gas-Sensorbox	
Plattform		Plattform	
Einplatinencomputer	Raspberry Pi 3 B	Einplatinencomputer	Raspberry Pi Zero W
Stromversorgung		Stromversorgung	
Netzteil	USB-Netzteil 2,5 A	Akku	USB-Powerbank 5,2 Ah
Sensoren		Geschätzte Laufzeit	18 Stunden
Multi-Gassensor	MiCS-6814 (I ² C)	Sensoren	
Feinstaubpartikelsensor	SDS011 (UART)	Multi-Gassensor	MiCS-6814 (I ² C)
Temperatur / Luftfeuchtigkeit	DHT22 (single wire)	Feinstaubpartikelsensor	SDS011 (UART)
MQ-Gassensoren	MQ2, MQ3, MQ5, MQ9 (AD-Wandler)	Temperatur / Luftfeuchtigkeit	SHT75 (two wire)
CO ₂ -Sensor	AirCO2ntrol Mini (USB)	Luftdruck / Temperatur	BMP280 (I ² C)
VOCs-Sensor	iAM (USB)	Sonstiges	Lüfter, RTC (I ² C)
Lichtsensor	BH1750 (I ² C)	Größe l × b × h	12,3 × 10,4 × 5,0 cm
Sonstiges	Lüfter, 16-bit-AD-Wandler (ADS1115)		
Größe l × b × h	14,4 × 8,4 × 6,7 cm		

TABELLE 3.4: Übersicht über die Komponenten der zwei stationären und der mobilen Gas-Sensorbox.

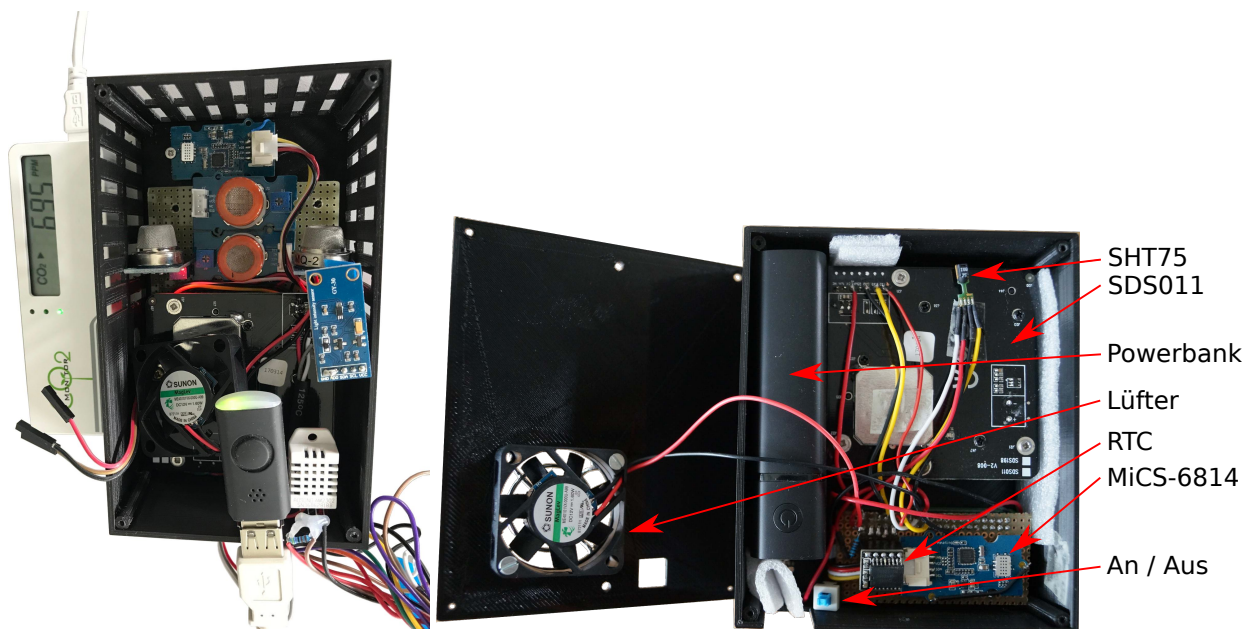


ABBILDUNG 3.16: Links: Eine der stationären Gas-Sensorboxen (Abbildung verändert aus (Dang, Seiderer & André, 2018)). Rechts: Mobile Gas-Sensorbox.

4 Echtzeitsignalverarbeitung und maschinelles Lernen

Bei der Verarbeitung von Sensordaten fallen zwei verschiedene Arten von Signalen an: kontinuierliche und diskrete. Ein Beispiel für ein kontinuierliches Signal ist ein Audiosignal, das mit einer konstanten Abtastrate (z. B. 48 kHz) erfasst wird. Diskrete oder ereignisbasierte Daten können z. B. von einem Bewegungsmelder stammen. Es ist allerdings auch möglich, die Signalarten in die jeweils andere umzuwandeln. Ereignisse können z. B. bei einem Audiosignal erzeugt werden, wenn ein Schwellwert überschritten wurde. Diese Ereignisse können wiederum mit einer spezifizierten Abtastrate in ein kontinuierliches Signal gewandelt werden, indem der letzte Wert wiederholt wird.

Abhängig von der jeweiligen Anwendung wurde in dieser Arbeit optimierte Software für kontinuierliche oder diskrete Signale eingesetzt, die in diesem Kapitel vorgestellt wird. Ergänzend werden wesentliche Verfahren des maschinellen Lernens kurz behandelt, die in verschiedenen eigenen Untersuchungen eingesetzt wurden.

4.1 Kontinuierliche Daten

Erfassung Kontinuierliche Daten werden meist mit einer möglichst stabilen Abtastrate erfasst und verarbeitet. Eventuelle Ungenauigkeiten bei der Abtastfrequenz müssen, je nach Verarbeitungsverfahren, kompensiert werden. Zu kontinuierlichen Datenquellen zählen beispielsweise Audiosignale oder auch viele Biosignale, die mit einem Analog-Digital-Wandler mit einer bestimmten Abtastrate und Quantisierung digitalisiert wurden.

Verarbeitung Bei kontinuierlichen Daten ist es für eine weitere Datenverarbeitung nötig, eine Segmentierung durchzuführen, wodurch sogenannte „Datenfenster“ (engl. data windows) entstehen. Auf diese Datenfenster können beispielsweise Filter aus der digitalen Signalverarbeitung (DSP) angewandt werden, die im Zeitbereich (z. B. IIR¹-Filter) oder im Frequenzbereich (z. B. FIR²-Filter) arbeiten.

Synchronisierung Eine ggf. erforderliche Datensynchronisation zwischen Signalströmen kann dadurch erreicht werden, dass die Abtastrate aller Datenströme auf eine gemeinsame Frequenz gesetzt wird. Dies ist mittels Abtastratenerhöhung (engl. upsampling) oder Heruntertaktung (engl. downsampling) möglich. Alternativ können bei einzelnen Samples auch Zeitstempel vorhanden sein, die zur Synchronisation verwendet werden können.

Fehlende Daten Generell besitzen alle Taktgeber, wie Oszillatoren, eine mehr oder weniger große Ungenauigkeit. Ungenauigkeiten bei der Taktfrequenz werden jedoch vor allem bei mobilen Geräten in Kauf genommen, da sie Energiesparmaßnahmen nutzen müssen. Mobile Betriebssysteme wie Android versuchen Hintergrundprogramme aggressiv anzuhalten oder zu beenden, um Sensorik oder CPU-Kerne möglichst häufig abschalten oder die Taktfrequenzen reduzieren zu können. Durch das Deaktivieren und Reaktivieren der Komponenten oder die Änderung der Taktraten

¹Infinite Impulse Response

²Finite Impulse Response

kommt es jedoch immer zu leichten Verzögerungen, die einen erhöhten Jitter (zeitliche Abweichungen von der Abtastrate) bewirken können (siehe hierzu die Sensor-Jitteranalyse für Android von Peguero et al. (2016)). Bei der Echtzeitsignalverarbeitung muss Jitter daher berücksichtigt werden, indem die Software mit fehlenden (missing data) oder zu vielen Werten umgehen kann.

Software Für die Echtzeitverarbeitung von kontinuierlichen Daten existieren diverse Softwarelösungen. Einige verfügen auch über Verfahren des maschinellen Lernens. Manche Lösungen sind ausschließlich auf bestimmte Datenarten spezialisiert, wie OpenSmile (Eyben et al., 2010) für Audio, GRT (Gillian & Paradiso, 2014) für Gestenerkennung mittels Beschleunigungsdaten oder OpenFace (Baltrušaitis et al., 2016) zur Erkennung von Gesichtsausdrücken mittels Bilderkennung. Es gibt zudem generische kommerzielle Produkte wie MathWorks MATLAB und Simulink. Ein neueres Framework ist das quelloffene MediaPipe von Google (Lugaresi et al., 2019), das sich allerdings noch im Alpha-Stadium befindet.

In dieser Arbeit wurden für viele Echtzeitanwendungen die am Lehrstuhl entwickelten, quelloffenen Tools SSI bzw. mobileSSI eingesetzt, die für beliebige Datenquellen und Einsatzzwecke nutzbar sind und im Nachfolgenden vorgestellt werden.

4.1.1 Das Social Signal Interpretation (SSI) Framework

Das SSI-Framework sowie dessen mobile Variante (mobileSSI) wurden in dieser Arbeit bei verschiedenen Projekten als Hilfsmittel eingesetzt. Vom Autor bzw. in Kooperation wurden hierfür diverse Erweiterungen entwickelt, die in Anhang C aufzufinden sind. Im folgenden Abschnitt wird das SSI-Framework vorgestellt. In der Publikation von Wagner et al. (2013) sowie auf der offiziellen Website³ wird die Software ausführlich behandelt.

Das Social Signal Interpretation (SSI) Framework⁴ ist eine quelloffene, in C++ geschriebene Tool-Sammlung für Windows PCs (x86) zur Echtzeitsignalverarbeitung, die vor allem auf Signale ausgelegt ist, die von Menschen stammen. Hierzu zählen die Stimme, Gestik und Mimik, die mit verschiedenen Sensoren in Echtzeit synchronisiert erfasst, segmentiert, verarbeitet und gespeichert werden können. Das Toolkit umfasst neben den Schnittstellen für verschiedene Sensoren und Komponenten zur Berechnung verschiedener Merkmale auch alle Schritte zur Erstellung und Evaluation von Klassifizierungsmodellen.

Für die Signalverarbeitung werden sogenannte Pipelines erstellt, die mittels C++-Code oder XML-Datei definiert werden können. Die Pipeline-Komponenten wurden als Plugins realisiert, die in C++ oder Python geschrieben werden können. Es existieren drei Arten von Pipeline-Komponenten: *Sensors*, *Transformers* und *Consumers*. Alle Komponenten besitzen hierbei einen eigenen Thread. Die Daten zwischen den Komponenten werden mit Ringpuffern ausgetauscht. In Abhängigkeit von den entsprechenden Komponenten besitzen diese eine definierte Dimension und festgelegten Datentyp. Sie können zudem Events (diskrete Daten) ausgeben und empfangen. Die Größe des Ringpuffers bzw. Datenfensters kann vom Nutzer definiert werden, hierbei ist es möglich, eine Überlappung anzugeben. Diese Datenfenster und Überlappungen müssen vom Nutzer für die darauffolgenden Pipeline-Komponenten sinnvoll gewählt werden.

Ein Beispiel einer einfachen Pipeline mit Eingabe und Ausgabe ist in Abbildung 4.1 dargestellt. Ein 3-Achsen-Beschleunigungssensor wird von einem Sensor-Plugin ausgelesen und liefert Gleitkommawerte (Datentyp *float*) mit einer Dimension von drei. Es sind daher bei diesem Beispiel nur *Transformer* oder *Consumer* nutzbar, die mit Gleitkommawerten und mehreren Dimensionen umgehen können. Als *Transformer* wäre daher die Berechnung verschiedener statistischer Merkmale (z. B. Maximum, Minimum, Varianz, Mittelwert, Korrelation) auf einer Dimension / Achse oder dimensionsübergreifend möglich. Der *Transformer* gibt daraufhin Merkmalsvektoren mit einer bestimmten Dimension (Merkmalsanzahl) aus, die anschließend in einen, bereits mit SSI trainierten,

³<http://openssi.net> (27.02.2020)

⁴<https://github.com/hcmlab/ssi> (27.02.2020)

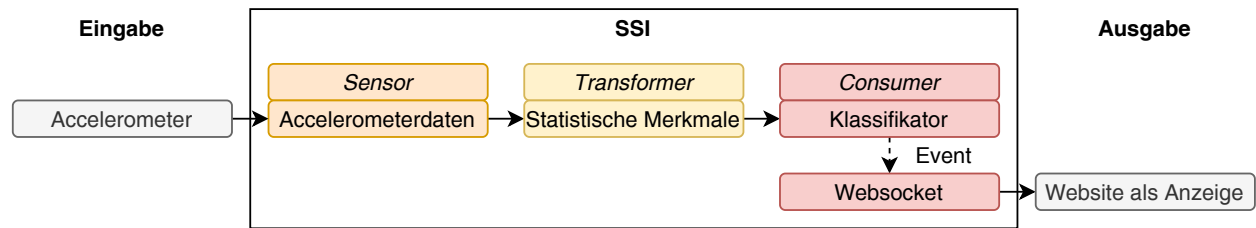


ABBILDUNG 4.1: Beispiel einer einfachen SSI-Pipeline.

Klassifikator (*Consumer*) geleitet werden können, um eine Echtzeitklassifizierung durchzuführen. Das Klassifikationsergebnis kann als Event intern versandt, vom WebSocket-Plugin (*Consumer*) empfangen, und schließlich auf einer Website dargestellt werden.

SSI setzt eine konstante Abtastrate bei Sensoren voraus, die mit einem eigenen Timer (Sleep-Funktion) forciert wird. Eine alternative Lösung der Synchronisation, die in SSI nicht verfügbar ist, hätte die Nutzung von Zeitstempeln dargestellt. Falls bestimmte Komponenten den festgelegten Takt nicht einhalten, werden bei SSI Mechanismen ausgelöst, fehlende Daten zu simulieren. Zum Beispiel durch Ergänzung von Null-Werten oder der Wiederholung alter Werte bzw. dem Auslassen überschüssiger Daten.

Fünf verschiedene Formen der Fusion von Informationen sind derzeit, laut der Dokumentation, in SSI bereits enthalten. Diese werden benötigt, um beispielsweise multimodale Klassifizierungen zu ermöglichen, bei denen die Daten meist von verschiedenen Sensoren mit unterschiedlichen Abtastraten, Datenfenstern (Zeitdauer / Überlappung) und Merkmalen stammen können.

Für maschinelles Lernen sind verschiedene Klassifikationsverfahren integriert. Hierzu zählen klassische Verfahren wie (lineare) Support Vector Machines und Naive-Bayes, aber auch aktuelle wie tiefe neuronale Netze.

4.1.2 MobileSSI

Bei mobileSSI⁵ handelt es sich um eine Portierung von SSI auf die Systeme Linux und Android (Flutura, Wagner, Lingenfeller, Seiderer & André, 2016c). Damit wird eine mobile Nutzung auf Smartphones und Smartwatches ermöglicht. Auch auf vielen ARM-basierten Geräten, wie Raspberry Pis, ist mobileSSI lauffähig. Ein einfach ausführbarer Docker-Container wurde hierfür vom Autor erstellt⁶. Viele der in SSI enthaltenen Plugins, die nicht spezielle Treiber oder plattformspezifische Schnittstellen bzw. Funktionen einsetzen, können auch unter Linux und Android genutzt werden.

Da Android über diverse Stromsparmechanismen und Berechtigungen verfügt, musste eine Android-App in Java implementiert werden, die mobileSSI-Pipelines als Hintergrunddienst mit den notwendigen Berechtigungen startet. Ein spezielles Plugin zum Auslesen von Sensordaten, die das Android-System verwaltet, wurde erstellt. Android verfügt allerdings nicht für alle Sensoren und die Bluetooth-Schnittstelle über C++-Schnittstellen, da häufig die Programmiersprache Java priorisiert wurde. Daher wurde eine optionale Anbindung an den SSI-Klon „SSJ“ (Damian et al., 2018), der ausschließlich für Android angedacht und in Java geschrieben ist, angebunden. Im Gegenzug erlaubt mobileSSI, SSJ die Nutzung vieler, bereits für SSI entwickelter, effizienter C++-Plugins, wie beispielsweise Filter oder einige Verfahren des maschinellen Lernens.

4.2 Diskrete / eventbasierte Daten

Ereignisbasierte (engl. eventbased) oder diskrete Sensordaten treten beispielsweise häufig im Smarthome-Bereich auf. Präzise Zeitintervalle sind hierbei aus folgenden Gründen nur selten

⁵<https://github.com/hcmlab/mobileSSI> (28.02.2020)

⁶<https://github.com/hcmlab/mobileSSI/wiki/mobileSSI-with-Docker%3A-run> (27.02.2020)

vorhanden:

- Um die Batterielaufzeit zu erhöhen und die Funkfrequenz weniger zu belasten, können Sensorgeräte bereits vor der Übertragung so eingestellt sein, dass nur veränderte Werte übertragen, ein bestimmter Schwellwert überschritten oder ein zeitlicher Mindestabstand zwischen dem letzten Sendevorgang erreicht werden muss.
- Die Auslastung verschiedener Frequenzbänder ist gesetzlich reguliert. In Deutschland ist hierfür die Bundesnetzagentur zuständig. Beim lizenzfreien 868 MHz-Band existieren beispielsweise maximale Frequenzbelegungsdauern, die nicht überschritten werden dürfen. Einige Smarthome-Systeme wie Homematic deaktivieren bei einer Überschreitung die Funkverbindung sofort vollständig. Maximale Sendeleistungen sind ebenfalls für verschiedene Frequenzen festgelegt und begrenzen deren Reichweite.
- Durch Übertragungsfehler können Datenpakete verloren gehen oder unkorrigierbar beschädigt werden. Beschädigte Pakete können beispielsweise mit der zyklischen Redundanzprüfung (CRC) oder Überprüfung der Datenparität erkannt werden, wenn in den jeweiligen Datenpaketen ein CRC-Wert bzw. Paritäts-Bit enthalten ist. Bei bidirektionalen Funkprotokollen werden Datenpakete, wenn sie beim Ziel ankommen, durch das Absenden einer Nachricht zur Quelle bestätigt (acknowledged). Falls die Quelle keine Bestätigung innerhalb eines bestimmten Zeitfensters erhält, sendet diese das ursprüngliche Datenpaket erneut.
- Unterschiedlich lange Latenzen können aus folgenden Gründen auftreten, wie z. B. Funkmodule warten auf ein freies Frequenzband, eine Datenweiterleitung erfolgt über mehrere Geräte (Mesh, Repeater), Datenverschlüsselung / -entschlüsselung, Fehlerkorrekturen oder rechenintensiver Datenverarbeitung.

Von mehreren Ereignissen, z. B. durch Aggregation abgeleitete Ereignisse, werden auch als komplexes Ereignis (engl. complex event) bezeichnet (Dayarathna & Perera, 2018). Verbreitete Frameworks wie z. B. Apache Spark (Zaharia et al., 2016) sind allerdings auf die Verarbeitung von „Big Data“ optimiert, also auf die Verarbeitung einer sehr großen Anzahl komplexer Ereignisse in kurzen Zeitintervallen. Entsprechende Frameworks besitzen daher vergleichsweise hohe Hardwareanforderungen. In einem privaten Smarthome treten allerdings wesentlich geringere Datenmengen auf, wodurch diese mit entsprechend optimierter Software auch von relativ leistungsschwachen, kostengünstigen Embedded Systemen, wie einem Raspberry Pi 3 mit nur geringem Energiebedarf, verarbeitet werden können. Die in dieser Arbeit entwickelten Systeme nutzten ebenfalls Embedded Hardware. Daher wurde ausschließlich ressourcenschonende Software eingesetzt und entwickelt.

4.2.1 Hausautomations-Software

Im Rahmen dieser Arbeit wurden ausschließlich Open-Source-Softwarelösungen genutzt, die Hardware diverser Hersteller und eigene Hardware ansteuern und beliebig angepasst werden können.

Eine Übersicht zu verbreiteter Open-Source-Hausautomations-Software ist in Tabelle 4.1 dargestellt. Sie wurde auch für die Auswahl geeigneter Softwarelösungen herangezogen, wobei zu beachten ist, dass erste Projekte in dieser Arbeit im Jahr 2014 implementiert wurden. Die generelle Verbreitung der entsprechenden Programme zu bestimmen, ist nur begrenzt möglich, da sie u. a. von der verfügbaren Hardware in einer Region abhängt auf die Entwickler und Nutzer zugreifen können. Auch die Popularität bestimmter Programmiersprachen spielt eine Rolle. In Internetsuchen und in Tutorials häufiger anzutreffende Programme sind ioBroker, OpenHAB, Home Assistant und FHEM. Die Anzahl an aktiven Nutzern lässt sich allerdings nicht exakt bestimmen. Als Anhaltspunkt können die Anzahl der registrierten Nutzer der offiziellen Foren dienen. Am 13.02.2021

waren dies, auf Tausender gerundet, bei Home Assistant ca. 89.000⁷, bei OpenHAB ca. 39.000⁸, bei ioBroker ca. 36.000⁹ und bei FHEM¹⁰ ca. 22.000 registrierte Nutzer.

Name	Program- miersprache	Jahr	Name	Program- miersprache	Jahr
MisterHouse	Perl	2000	Home Assistant	Python	2013
FHEM	Perl	2005	Gladys	Javascript	2013
DomotiGa	Gambas	2008	PiDome	Java	2013
domogik	Python	2008	ago control	C++	2014
OpenHAB	Java	2010	HomeGenie	C#	2014
HoMIDoM	VB	2011	ioBroker	Javascript	2014
Domoticz	C++	2012	Pimatic	Javascript	2014
OpenMotics	Python	2012	Jeedom	PHP	2014
Freedomotic	Java	2012	MyController.org	Java	2015
Pytotation	Python	2012	Homify	PHP	2017
MajorDoMo	PHP	2012	NimHA	Nim	2018
Calaos	C++	2013			

TABELLE 4.1: Übersicht über Hausautomations-Software sortiert nach Jahr der Veröffentlichung (Datum des ersten Code-Uploads, Copyright oder Angabe auf Website), die per Internetsuche gefunden werden konnte. Relevantere Software ist fett hervorgehoben. Links zur entsprechenden Website oder Github-Projekt sind in der PDF-Version enthalten.

In dieser Arbeit wurde in mehreren Projekten FHEM eingesetzt, da es unter Windows und Linux bei eigenen Langzeittests stabil lief und zudem beispielsweise im Vergleich zu OpenHAB sehr geringe Ansprüche an die Hardware stellt. Zusätzlich besitzt es eine sehr gute Unterstützung von Homematic-Hardware, die insbesondere in Deutschland verbreitet ist. In mehreren Untersuchungen wurde die Software auf Raspberry Pis betrieben.

4.2.2 Eventerpretor

Für die Verarbeitung von Ereignissen und für die Kontextinterpretation wurde bei einigen Untersuchungen (siehe Abschnitt 7.3.4 und 8.3.6) die eigene quelloffene Software „Eventerpretor“ verwendet, deren Quelltext auf Github verfügbar ist¹¹. Die Software ist in Kotlin geschrieben (kompatibel mit Java) und wurde unter Linux und Windows eingesetzt. Die Software kann zudem unter Android in einer Chroot-Umgebung verwendet werden. Die Software ist im Gegensatz zu (mobile)SSI auf die Verarbeitung von Ereignissen optimiert und wird ausschließlich aktiv, wenn neue Ereignisse eintreffen. Hierdurch spart sie CPU-Leistung. Eine höhere Dateninterpretation mit eigenem Sensormodell bei der ein Regelsystem (Drools¹²) genutzt werden kann, ist möglich. In Listing 4.1 ist eine Regel dargestellt, die die Luftqualität des CO₂-Gassensors „AirCO2ntrol“ mittels gut, neutral und schlecht interpretiert und ein entsprechendes Kontextelement anlegt, das von anderen Regeln genutzt werden kann.

Eine effiziente Nutzung von PostgreSQL (u. a. mit asynchronen Nachrichten und Triggerfunktionen) für datenbankbasierte Aggregationen über längere Zeitbereiche ist integriert, wodurch komplexe Operationen direkt mittels Datenbank durchgeführt werden können. Es wird eine flow-basierte Programmierung mit wiederverwendbaren Knoten (nodes) eingesetzt. Jeder Knoten nutzt

⁷<https://community.home-assistant.io/about> (13.02.2021)

⁸<https://community.openhab.org/u?period=all> (13.02.2021)

⁹<https://forum.iobroker.net/category/4/deutsch> (13.02.2021)

¹⁰<https://forum.fhem.de/> (13.02.2021)

¹¹<https://github.com/andreas-seiderer/Eventerpretor>

¹²<https://www.drools.org/> (18.09.2019)

hierbei einen eigenen Thread. Die Kommunikation zwischen den Knoten findet threadsafe mittels *LinkedBlockingQueues* statt.

```
rule AirQuality when
  s: Sensor(sensorname=="AirCO2ntrol" && sensortype=="gas_sensor" && value!=null)
  if((double)s.getValue() < 1000.0) do[good]
  if((double)s.getValue() >= 1000.0 && (double)s.getValue() < 2000.0) do[neutral]
  if((double)s.getValue() >= 2000.0) do[bad]
then
then[good]
  insertLogical(new ContextElement("good", "AirQuality", s));
then[neutral]
  insertLogical(new ContextElement("neutral", "AirQuality", s));
then[bad]
  insertLogical(new ContextElement("bad", "AirQuality", s));
end
```

LISTING 4.1: Drools-Regel zur Interpretation der Luftqualität.

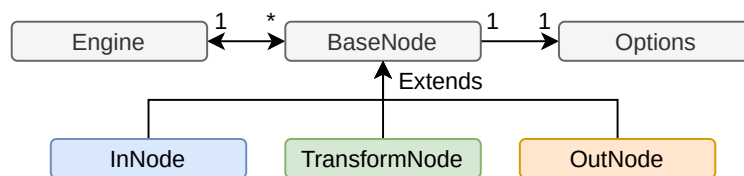


ABBILDUNG 4.2: Vereinfachtes Klassendiagramm von Eventerpretor.

In Abbildung 4.2 ist ein vereinfachtes Klassendiagramm von Eventerpretor abgebildet. Bei der Ausführung wird eine Instanz der *Engine* erzeugt, die beliebig viele *InNodes*, *TransformNodes* und *OutNodes* verwaltet. Diese Knoten erweitern die Klasse *BaseNode*. Jeder Knoten besitzt hierbei spezifische Optionen, die als JSON-Dateien gespeichert werden. Alle implementierten Knoten erben von diesen drei Basisknoten. *InNodes*, wie beispielsweise ein Knoten, der MQTT-Nachrichten empfängt, geben ausschließlich Daten aus. *TransformNodes* verarbeiten Daten und besitzen einen Eingang und Ausgang. Hierzu zählt beispielsweise das Regelsystem. *OutNodes* besitzen ausschließlich einen Eingang. Hierzu zählt z. B. ein Knoten, der Daten in eine PostgreSQL-Datenbank einfügt.

Aus den Knoten können Pipelines erzeugt werden, indem entsprechende Knoten miteinander verbunden werden (z. B. `node1.publishTo(node2)`). Die Definition von Pipelines findet in Groovy-Skripten statt. Eine Instanz von Eventerpretor kann mehrere Pipelines ausführen. Die Knoten können beliebig miteinander verknüpft werden, es muss allerdings teilweise auf kompatible Datentypen geachtet werden. Zudem findet keine Überprüfung auf Zyklen statt. Eine GUI (Website) mit grafischer Darstellung der Knotenverbindungen ist vorhanden und ermöglicht das Auffinden von Verbindungsfehlern. Die GUI erlaubt zusätzlich ein sauberes Herunterfahren der Pipelines sowie ein Zurücksetzen einzelner Knoten während der Laufzeit, um neue Optionen zu laden. Für manuelle Dateneingaben können ebenfalls Webinterfaces genutzt werden, um beispielsweise Tests und Simulationen zu ermöglichen. Ein ausführlicheres Beispiel mit mehreren Pipelines und Simulations-Websites mit Abbildungen ist in Abschnitt 7.3.4 aufzufinden.

Eventerpretor besitzt einige Ähnlichkeiten zu IBMs Node-RED¹³, Eventerpretor wurde allerdings explizit für die Anforderungen bei den Projekten angepasst. Durch die Verwendung von Java statt Javascript mit Node.js verfügt es über Vorteile bezüglich der Parallelisierung und Nutzung von vorhandenen Java-Programmbibliotheken.

Eventerpretor besitzt einige Knoten, die speziell für die Erweiterung von FHEM ausgelegt sind. Hierdurch kann beispielsweise das Sensormodell und Regelsystem von Eventerpretor mit Daten von FHEM eingesetzt werden und Befehle zurück an FHEM schicken. Im Gegensatz zu FHEM

¹³<https://nodered.org/> (18.09.2019)

arbeitet Eventerpretor hierbei multithreaded und verringert somit Latenzen bei komplexeren Berechnungen. Dies ist insbesondere beim Einsatz mit Raspberry-Pi-3B-Einplatinencomputern von Vorteil, die eine geringe Singlecore-CPU-Leistung besitzen.

4.3 Maschinelles Lernen

Dieser Abschnitt enthält einen Überblick zu den primär in dieser Arbeit eingesetzten Verfahren des maschinellen Lernens und der verwendeten deutschen Terminologie.

4.3.1 Eingesetzte Verfahren

Reinforcement Learning

Reinforcement Learning oder auch bestärkendes Lernen wird im Buch von Sutton und Barto (1998) ausführlich vorgestellt. Es wurde in den Abschnitten 8.3.6 und 9.7 für Anpassungen der sprachlichen Formulierungen genutzt.

Naive-Bayes

Der Naive-Bayes-Algorithmus kann für Klassifizierungsprobleme beim überwachten maschinellen Lernen eingesetzt werden und basiert auf dem Bayes-Theorem. Der grundsätzliche Algorithmus wurde von Knuth (1985, S. 115) beschrieben. Ein Naive-Bayes-Klassifikator, der so angepasst wurde, dass er für Inkrementelles- oder Online-Lernen genutzt werden kann, wurde in Abschnitt 8.3.5 eingesetzt.

Entscheidungsbäume

Entscheidungsbäume (engl. decision trees) gehören zu den überwachten maschinellen Lernverfahren. In den Abschnitten 8.3.2 und 8.3.3 wurde der C4.5-Algorithmus (Quinlan, 1993) zur Erstellung von Entscheidungsbäumen zur Klassifizierung angewandt.

Support Vector Machine (SVM)

Support Vector Machines wurden von Cortes und Vapnik (1995) präsentiert. Sie zählen zum überwachten maschinellen Lernen und können zur Klassifikation und Regression eingesetzt werden. Hierbei wird eine Hyperebene erzeugt, die Daten voneinander trennt. Es können verschiedene „Kernel“ genutzt werden, wobei eine lineare Funktion, den einfachsten Kernel darstellt. In dieser Arbeit wurden sie in folgenden Abschnitten für Klassifizierungen eingesetzt: 6.3.2, 6.3.1, 7.3.2, 8.3.1, 8.3.5. In Abschnitt 8.3.3 wurden sie für die Klassifikation und Regression (Vorhersage einer abhängigen Variable mit einer unabhängigen) verwendet.

Tiefe neuronale Netze

In dieser Arbeit wurden auch moderne tiefe neuronale Netze (engl. deep neural networks) eingesetzt, die Teil des sogenannten „Deep Learnings“ sind. Ein umfassendes, aktuelles Buch wurde hierzu von Goodfellow et al. (2016) verfasst. In folgenden Abschnitten wurden verschiedene Architekturen für Klassifizierungen genutzt: 7.3.1, 7.3.2 und 5.5.2. Die eingesetzte Spracherkennung und Sprachverarbeitung in Abschnitt 8.3.7 basierte ebenfalls auf tiefen neuronalen Netzen.

Tiefe neuronale Netze benötigen große Datenmengen, die allerdings nicht immer zur Verfügung stehen. Zudem ist das Training trotz moderner Frameworks mit GPU- oder TPU-Beschleunigung sehr zeitaufwändig. In solchen Fällen kann Transferlernen eine gute Lösung darstellen, bei der bereits trainierte Netze für eine ähnliche Aufgabe angepasst werden können. Grundlagen zum

Transferlernen werden in Kapitel 5 behandelt. Ein weiteres Problem stellt die Erstellung optimaler Netzarchitekturen dar. Hierzu wurde in Abschnitt 7.3.1 ein genetischer Algorithmus untersucht, der automatisch praktikable, optimierte Architekturen erstellen kann.

4.3.2 Klassifizierung

In diesem vereinfachten Abschnitt zu Klassifizierung werden die in dieser Arbeit verwendeten Techniken mit ihren deutschen Begriffen grundsätzlich eingeführt. Bei der Klassifizierung entscheidet ein Klassifizierungsmodell zu welcher Klasse Daten gehören. Hierbei können Daten nur den Klassen zugeordnet werden, die dem Modell bekannt sind. Für das Training eines Modells wird beim überwachten maschinellen Lernen ein automatisch oder manuell annotierter bzw. beschrifteter (engl. labeled) Datensatz eingesetzt. Die Datenbeschriftungen von Datenbereichen entsprechen meist den Klassen, die ein Klassifikator unterscheiden soll.

Datenkorpus Falls bei einem Datenkorpus die Anzahl der Daten pro Klasse nicht ausgeglichen (engl. balanced) ist, wird dies als unausgeglichener (engl. imbalanced) Datensatz bezeichnet. Je nach Art des Klassifikators kann dies zu einer Gewichtung der Klassen führen, sodass Klassen mit einer geringeren Datenmenge schlechter erkannt werden als die mit einer größeren Menge. Für dieses Problem existieren verschiedene Methoden, um die Datenmengen pro Klasse anzugleichen. Ein solches Verfahren ist das „random under-sampling“. Hierbei wird zunächst die Datenanzahl der Klasse mit der geringsten Datenmenge bestimmt. Anschließend werden bei allen anderen Klassen so lange Daten zufällig entfernt, bis deren Datenmenge der Datenanzahl der „kleinsten“ Klasse entspricht.

Datensegmentierung und Merkmalsextraktion Je nach Art der Daten und dem verwendeten Klassifizierungsverfahren ist eine Segmentierung in Datenfenster mit einer bestimmten Größe notwendig. Bei kontinuierlichen Daten, wie z. B. Audiodaten, sind dies Zeitbereiche. Die Größe dieser Bereiche beeinflusst hierbei die Latenz einer Echtzeitklassifizierung, da dieser Puffer zunächst mit Daten befüllt werden muss. Eine Überlappung von Datenfenstern ist möglich und kann bei bestimmten Anwendungen vorteilhaft sein. Basierend auf den Werten von Datenfenstern werden in vielen Fällen Merkmale (engl. features) berechnet. Dieser Vorgang wird auch als Merkmalsextraktion (engl. feature extraction) bezeichnet. Merkmale sind hierbei häufig statistische Funktionen, wie das Minimum oder der Mittelwert der Werte des Datenfensters. Bei Audiosignalen werden beispielsweise in vielen Fällen zusätzlich frequenzbasierte Merkmale eingesetzt. Auch neuronale Netze können zur Erzeugung von Merkmalen verwendet werden. Dies wird z. B. beim Transferlernen (siehe Kapitel 5) genutzt. Die einzelnen Merkmale werden in einem Merkmalsvektor gespeichert und werden für das Training eines Klassifizierungsmodells eingesetzt.

Training und Evaluation von Modellen Beim Training und der Evaluation eines Klassifikators mit einer festen Aufteilung des Datensatzes wird dieser mindestens in die disjunkten Teilmengen, Trainings- und Test-Daten, aufgeteilt (engl. split). Beim Training neuronaler Netze wird das Datenkorpus häufig zusätzlich in eine Validierungs- (engl. validation dataset) oder Entwicklungs-Datenmenge (engl. development dataset) unterteilt. Diese Validierungs-Daten werden zur Evaluation der Leistung des Modells während des Trainings mit den Trainings-Daten genutzt. Die Test-Daten werden generell verwendet, um das finale Modell zu evaluieren. Diese Aufteilung eines Datenkorpus dient dazu, die Leistungsfähigkeit eines Modells besser abschätzen zu können, da das Modell auf die Test-Daten nicht optimiert wurde.

Bei einer k-fachen Kreuzvalidierung (engl. k-fold cross-validation) werden k Unterteilungen des Datensatzes in disjunkte Trainings- und Test-Datenmengen für das Training und die Evaluation der Modelle durchgeführt. Die Datenpartitionierung kann hierbei k-fach zufällig durchgeführt werden. Eine spezielle Form der Kreuzvalidierung ist die Leave-One-Out-Kreuzvalidierung (engl.

leave-one-out cross-validation). Hierbei werden beispielsweise die Daten einer Person aus dem Trainings-Datensatz herausgenommen und stattdessen als Test-Daten eingesetzt. Dies wird für alle Personen durchgeführt. Da es sich in der eigenen Arbeit in Abschnitt 8.3.1 um verschiedene Sprecher handelte, wurde diese Evaluationsmethode Leave-One-Speaker-Out genannt.

In dieser Arbeit werden die deutschen Begriffe (englisch in Klammern) für die folgenden Gütemaße bei der Evaluation von Klassifikatoren verwendet (TP = true positive, TN = true negative, FP = false positive, FN = false negative):

$$\text{Treffergenauigkeit}(\text{accuracy}) = \frac{TP + TN}{TP + TN + FP + FN} \quad (4.1)$$

$$\text{Trefferquote}(\text{recall}) = \frac{TP}{TP + FN} \quad (4.2)$$

$$\text{Genauigkeit}(\text{precision}) = \frac{TP}{TP + FP} \quad (4.3)$$

5 Transferlernen

Beim Transferlernen (engl. transfer learning) wird bereits erlerntes Wissen genutzt, um neue Probleme schneller lösen zu können oder bessere Lösungsmöglichkeiten zu finden (Pan & Yang, 2010). Bezogen auf den Menschen und die Domäne „Programmiersprachen“ kann beispielsweise das Erlernen von C hilfreich sein, um z. B. andere imperative Programmiersprachen wie C++ oder auch Java später leichter lernen zu können, da bereits grundsätzliche Prinzipien des Programmierens bekannt sind.

Bei maschinellem Lernen verhält es sich ähnlich. Im Gegensatz zum klassischen Ansatz, bei dem für jeden neuen Datensatz ein komplett neues Modell erzeugt wird, werden bereits vorhandene Modelle genutzt, um ein neues Modell mit weniger Ressourcen oder auch Trainings-Daten erzeugen zu können oder dessen Performanz zu verbessern. Insbesondere bei rechen- und datenintensiven Verfahren, wie tiefen neuronalen Netzen, spielt Transferlernen eine bedeutende Rolle. Hierdurch ist es beispielsweise möglich, auch auf nutzereigenen mobilen Geräten und Embedded Systemen tiefe neuronale Netze zu trainieren und somit für nutzerspezifische Aufgaben anzupassen. Dem entsprechend kann die Privatsphäre des Nutzers geschützt werden, da keine persönlichen Daten an externe Cloud-Dienste übertragen werden müssen. Das grundsätzliche Konzept des Transferlernens wird bereits seit längerer Zeit in der Forschung verfolgt. Die folgenden Forschungsgebiete sind laut Pan und Yang (2010) mit dem Transferlernen verwandt oder besitzen vergleichbare Ziele:

- learning to learn
- inductive transfer
- knowledge transfer
- metalearning
- life-long learning
- multitask learning
- knowledge consolidation
- context-sensitive learning
- knowledge-based inductive bias
- incremental / cumulative learning

Wie in Abbildung 5.1 ersichtlich ist, hat in den letzten Jahren, mit dem Aufkommen von Deep Learning, die Bedeutung von Transferlernen deutlich zugenommen.

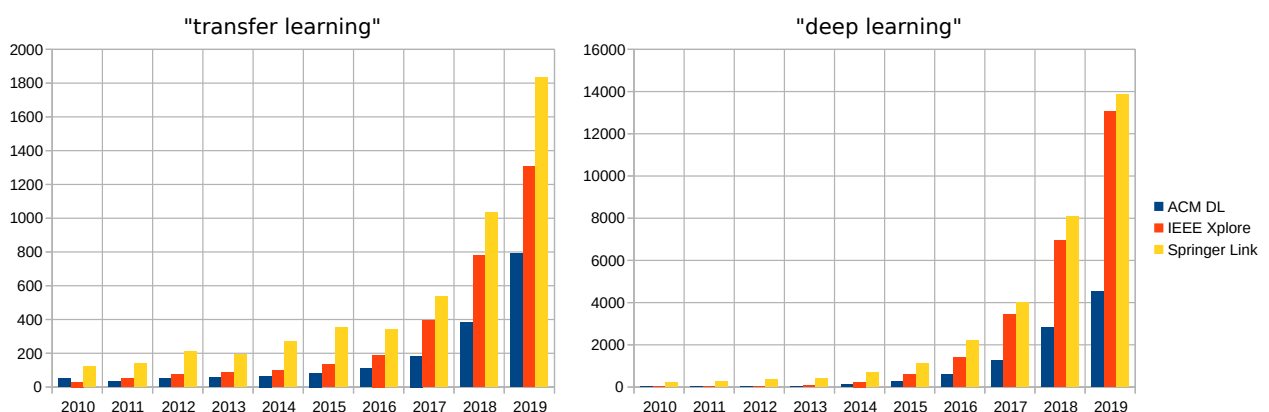


ABBILDUNG 5.1: Anzahl der Publikationen pro Jahr (2010-2019) in drei, in der Informatik relevanten, digitalen Bibliotheken. Die Art der Publikation wurde nicht eingeschränkt. Duplikate zwischen den Datenbanken sind vereinzelt möglich. Die Daten wurden am 16.10.2020 erhoben.

Zunächst werden in den folgenden Abschnitten grundsätzliche Definitionen eingeführt sowie ein Überblick darüber gegeben, wie Transferlernen kategorisiert werden kann. Anschließend wird die Anwendung bei Deep Learning genauer mit eigenen Anwendungen behandelt.

5.1 Definitionen für Domäne, Aufgabe und Transferlernen

Die folgenden formalen Definitionen stammen aus der Pionierarbeit von Pan und Yang (2010) und wurden hier sinngemäß übersetzt und zusammengefasst. Diese Definitionen werden in aktueller Literatur z. B. in den Arbeiten von Tan et al. (2018) und K. Weiss et al. (2016) und auch in der vorliegenden Arbeit genutzt.

Definition 5.1.1: Domäne (Pan & Yang, 2010)

Eine Domäne \mathcal{D} besteht aus einem Merkmalsraum^a \mathcal{X} und einer Randverteilung $P(X)$, wobei $X = \{x_1, \dots, x_n\} \in \mathcal{X}$. x_i ist der *i*-te Merkmalsvektor in einer Lernstichprobe^b X . Eine Domäne wird folgendermaßen definiert: $\mathcal{D} = \{\mathcal{X}, P(X)\}$.

^aengl. feature space

^bengl. learning sample

Definition 5.1.2: Aufgabe (Pan & Yang, 2010)

Eine Aufgabe \mathcal{T} besteht aus einem Beschriftungsraum^a \mathcal{Y} und einer Zielfunktion $f(\cdot)$, die nicht beobachtet, aber mit den Trainings-Daten gelernt werden kann. $\mathcal{T} = \{\mathcal{Y}, f(\cdot)\}$. Die Trainings-Daten bestehen aus Wertepaaren $\{x_i, y_i\}$, wobei $x_i \in \mathcal{X}$ und $y_i \in \mathcal{Y}$. Die Funktion $f(\cdot)$ kann zum Vorhersagen der entsprechenden Beschriftung genutzt werden und kann daher auch als eine bedingte Wahrscheinlichkeit $P(y|x)$ betrachtet werden. Daher wird eine Aufgabe als $\mathcal{T} = \{\mathcal{Y}, P(Y|X)\}$ definiert.

^aengl. label space

Definition 5.1.3: Daten Quell- und Zieldomäne (Pan & Yang, 2010)

- Daten der Quelldomäne^a: $D_S = \{(x_{S1}, y_{S1}), \dots, (x_{Sn}, y_{Sn})\}$, $x_{Si} \in \mathcal{X}_S$, $y_{Si} \in \mathcal{Y}_S$
- Daten der Zieldomäne^b: $D_T = \{(x_{T1}, y_{T1}), \dots, (x_{Tn}, y_{Tn})\}$, $x_{Ti} \in \mathcal{X}_T$, $y_{Ti} \in \mathcal{Y}_T$

In den meisten Fällen gilt für die Datenanzahl der Quell- bzw. Zieldomäne: $0 \leq n_T \ll n_S$

^aengl. source domain

^bengl. target domain

Definition 5.1.4: Transferlernen (Pan & Yang, 2010)

Sind eine Quelldomäne \mathcal{D}_S und eine Lernaufgabe^a \mathcal{T}_S , eine Zieldomäne \mathcal{D}_T und eine Lernaufgabe \mathcal{T}_T gegeben, dann hilft Transferlernen dabei das Lernen der Zielfunktion $f_T(\cdot)$ in \mathcal{D}_T zu verbessern, indem das Wissen aus \mathcal{D}_S und \mathcal{T}_S genutzt wird. Hierbei ist $\mathcal{D}_S \neq \mathcal{D}_T$ und / oder $\mathcal{T}_S \neq \mathcal{T}_T$.

^aengl. learning task

Für **traditionelles maschinelles Lernen** gilt, dass die Quell- und Zieldomäne sowie deren Lernaufgaben identisch sind: $\mathcal{D}_S = \mathcal{D}_T$ und $\mathcal{T}_S = \mathcal{T}_T$.

5.2 Probleme beim Transferlernen

Beim Transferlernen werden laut Pan und Yang (2010) folgende drei Forschungsprobleme untersucht, die auch bei der Anwendung behandelt werden müssen:

Welches Wissen wird transferiert? Welcher Teil vom Wissen kann zwischen Domänen oder Aufgaben übertragen werden?

Wie wird das Wissen transferiert? Mit welchem Algorithmus kann bestimmtes Wissen zum Lernen übertragen werden?

Wann sollte Wissen transferiert werden? In welchen Situationen kann Wissen übertragen werden bzw. wann sollte es unterlassen werden? Nicht immer kann Wissen mit Nutzen übernommen werden. Falls beispielsweise die Quell- und Zieldomäne keinen Bezug zueinander besitzen, kann ein Wissensübertrag in die Zieldomäne das Lernen verlangsamen oder sogar unmöglich machen.

5.3 Kategorien von Transferlernen

Die Kategorisierung von Transferlernen und deren Beschreibungen aus der Arbeit von Pan und Yang (2010) werden im folgenden Abschnitt sinngemäß wiedergegeben. Ein Überblick ist in Abbildung 5.2 dargestellt.

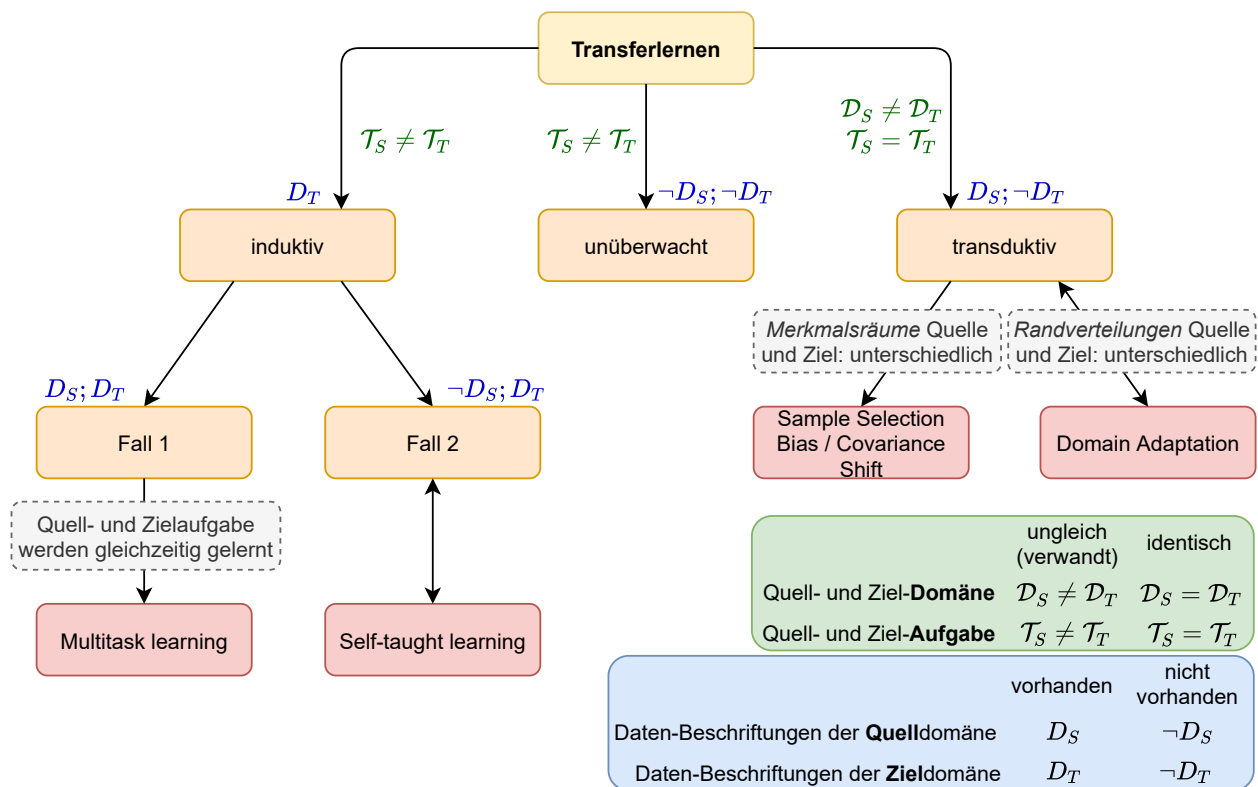


ABBILDUNG 5.2: Überblick über die Kategorisierung von Transferlernen mittels der Relation zwischen Quell- und Zieldomänen / Aufgaben sowie der Datenverfügbarkeit. Die Beziehung zu ähnlichen / identischen Verfahren sind ebenfalls ersichtlich (erweiterte / veränderte Zusammenfassung nach Pan & Yang, 2010).

5.3.1 Induktives Transferlernen

Beim induktiven Transferlernen können die Quell- und Zieldomäne gleich oder ungleich sein; die Quell- und Zielaufgabe unterscheiden sich, haben jedoch einen Bezug zueinander. In der Zieldomäne werden zum Lernen beschriftete Daten benötigt. Je nach Verfügbarkeit von beschrifteten Daten in der Quelldomäne können zwei Fälle unterschieden werden.

Fall 1 Es sind **viele beschriftete Daten** in der Quelldomäne vorhanden. In diesem Fall ist das induktive Transferlernen ähnlich zum Multitask-Lernen (engl. multitask learning). Beim induktiven Transferlernen wird ausschließlich versucht, die Zielaufgabe möglichst gut zu lernen, beim Multitask-Lernen werden Ziel- und Quellaufgabe gleichzeitig gelernt.

Fall 2 Es sind **keine beschrifteten Daten** in der Quelldomäne verfügbar. In diesem Fall ist das induktive Transferlernen vergleichbar mit dem autodidaktischen Lernen (engl. self-taught learning). Der Beschriftungsraum zwischen Quell- und Zieldomäne kann beim autodidaktischen Lernen unterschiedlich sein, sodass wie beim induktiven Transferlernen keine beschrifteten Daten in der Quelldomäne nutzbar sind.

Definition 5.3.1: Induktives Transferlernen (Pan & Yang, 2010)

Sind eine Quelldomäne \mathcal{D}_S und eine Lernaufgabe \mathcal{T}_S , eine Zieldomäne \mathcal{D}_T und eine Lernaufgabe \mathcal{T}_T gegeben, dann zielt induktives Transferlernen darauf ab, das Lernen der Zielfunktion des Ziels $f_T(\cdot)$ in \mathcal{D}_T mit dem Wissen aus \mathcal{D}_S und \mathcal{T}_S zu verbessern, wobei $\mathcal{T}_S \neq \mathcal{T}_T$.

5.3.2 Unüberwachtes Transferlernen

Wie beim induktiven Transferlernen unterscheiden sich beim unüberwachten Transferlernen (engl. unsupervised transfer learning) Quell- und Zielaufgabe, sie müssen allerdings auch hier miteinander verwandt sein. Es sind keine beschrifteten Daten vorhanden. Das unüberwachte Transferlernen wird vor allem für unüberwachtes maschinelles Lernen, wie z. B. Clustering, in der Zieldomäne genutzt.

Definition 5.3.2: Unüberwachtes Transferlernen (Pan & Yang, 2010)

Sind eine Quelldomäne \mathcal{D}_S und eine Lernaufgabe \mathcal{T}_S , eine Zieldomäne \mathcal{D}_T und eine Lernaufgabe \mathcal{T}_T gegeben, dann zielt unüberwachtes Transferlernen darauf ab, das Lernen der Zielfunktion des Ziels $f_T(\cdot)$ in \mathcal{D}_T mit dem Wissen aus \mathcal{D}_S und \mathcal{T}_S zu verbessern, wobei $\mathcal{T}_S \neq \mathcal{T}_T$ und \mathcal{Y}_S und \mathcal{Y}_T sind nicht beobachtbar.

5.3.3 Transduktives Transferlernen

Generell gilt beim transduktiven Lernen, dass im Gegensatz zum induktiven Lernen kein generisches Modell erzeugt wird. Es wird ausschließlich versucht, die Beschriftungen für vorhandene, unbeschriftete Daten vorherzusagen (Vapnik, 1998, S. 12f.). Für das transduktive Transferlernen wird hier die Bedingung, dass alle unbeschrifteten Daten in der Zieldomäne für das Training vorhanden sein müssen, gelockert, indem in diesem Fall nur ein Teil der unbeschrifteten Daten beim Training verfügbar sein müssen.

Beim transduktiven Transferlernen sind die Quell- und Zielaufgaben identisch, die Quell- und Zieldomänen unterscheiden sich. Es sind viele Daten in der Quelldomäne, jedoch keine Daten in der Zieldomäne vorhanden. Das transduktive Transferlernen kann in zwei weitere Unterkategorien unterteilt werden, indem der Unterschied zwischen den Domänen (Erinnerung: $\mathcal{D} = \{\mathcal{X}, P(X)\}$; siehe Definition 5.1.1) betrachtet wird.

Fall 1 Die Domänen unterscheiden sich nur bezüglich ihres Merkmalsraumes: $\mathcal{X}_S \neq \mathcal{X}_T$. In diesem Fall ist transduktives Transferlernen vergleichbar mit einer Stichprobenverzerrung (engl. sample selection bias) und kovarianten Verschiebung (engl. covariate shift).

Fall 2 Die Domänen unterscheiden sich nur bezüglich ihrer Randverteilung: $P(X_S) \neq P(X_T)$. In diesem Fall ist transduktives Transferlernen verwandt mit der Domänenanpassung (engl. domain adaptation).

Definition 5.3.3: Transduktives Transferlernen (Pan & Yang, 2010)

Sind eine Quelldomäne \mathcal{D}_S und eine Lernaufgabe \mathcal{T}_S , eine Zieldomäne \mathcal{D}_T und eine Lernaufgabe \mathcal{T}_T gegeben, dann zielt transduktives Transferlernen darauf ab, das Lernen der Zielfunktion des Ziels $f_T(\cdot)$ in \mathcal{D}_T mit dem Wissen aus \mathcal{D}_S und \mathcal{T}_S zu verbessern, wobei $\mathcal{D}_S \neq \mathcal{D}_T$ und $\mathcal{T}_S = \mathcal{T}_T$. Zusätzlich müssen während des Trainings unbeschriftete Daten aus der Zieldomäne vorhanden sein.

5.4 Ansätze beim Transferlernen

Um Transferlernen durchführen zu können, existieren verschiedene Ansätze, die von Pan und Yang (2010) zusammengestellt wurden:

Wissenstransfer von Instanzen Beim Wissenstransfer von Instanzen können beschriftete Daten aus der Quelldomäne durch Neugewichtung genutzt werden, um sie in der Zieldomäne beim Trainieren verwenden zu können.

Wissenstransfer von Merkmalsrepräsentationen Beim Wissenstransfer mittels Merkmalsrepräsentationen wird versucht, Merkmale zu finden, die die Unterschiede zwischen der Quell- und Zieldomäne reduzieren, sodass der Fehler von Klassifizierungs- und Regressionsmodellen verringert wird. Die Erzeugung von Merkmalsrepräsentationen kann hierbei überwacht oder unüberwacht stattfinden.

Wissenstransfer von Parametern Beim Wissenstransfer mittels Parametern können bestimmte Parameter oder A-priori-Wahrscheinlichkeiten von Hyperparametern genutzt werden, die von der Quell- und Zieldomäne geteilt werden.

Wissenstransfer von relationalem Wissen Der Wissenstransfer von relationalem Wissen wird bei Daten genutzt, die durch mehrere Relationen repräsentiert werden können. Daher geht dieser Ansatz nicht davon aus, dass die Daten unabhängig und identisch verteilt sind.

5.5 Transferlernen beim Deep Learning

In diesem Kapitel wurden bisher mittels des Literaturüberblicks von Pan und Yang (2010) grundsätzliche Definitionen für das Transferlernen beim maschinellen Lernen eingeführt sowie aufgezeigt, wie es eingeordnet werden kann. In diesem Abschnitt wird die Anwendung von Transferlernen für Deep-Learning-Szenarien genauer beleuchtet.

Beim Deep Learning werden große Datenmengen benötigt, um leistungsfähige Modelle trainieren zu können. Dieser Vorgang benötigt bei großen neuronalen Netzen und großen Datenmengen viel Rechenzeit, Arbeitsspeicher und Festplattenspeicher. In einigen Fällen ist es zudem nicht oder nur begrenzt möglich, mehr Daten zu gewinnen. Dies gilt beispielsweise für medizinische Bilder von seltenen Krankheiten, bei denen aufwändige Laboruntersuchungen notwendig sein können, um diese korrekt zu beschriften. Diese Probleme können mittels Transferlernen reduziert werden.

Nach der Definition von Tan et al. (2018) handelt es sich um tiefes Transferlernen (engl. deep transfer learning), falls beim Transferlernen (siehe Definition 5.1.4) die Zielfunktion $f_T(\cdot)$ eine nicht

lineare Funktion ist, die durch ein tiefes neuronales Netz ausgedrückt wird. In ihrer Arbeit haben Tan et al. (2018) folgende Techniken beim tiefen Transferlernen identifiziert:

Wissenstransfer von Instanzen Dieser Ansatz ist identisch zu dem bereits in Abschnitt 5.4 vorgestellten.

Wissenstransfer durch Abbildung Instanzen aus zwei Domänen werden in einen neuen Datenraum mit einer höheren Ähnlichkeit abgebildet.

Wissenstransfer mittels eines Adversarial-Ansatzes Nutzung eines Adversarial-Ansatzes („adversarial“ entspr. „gegnerisch“), um übertragbare Merkmale zu finden, die für beide Domänen nutzbar sind.

Wissenstransfer mittels Netzwerk Ein Teil eines trainierten Netzes aus der Quelldomäne wird weiterverwendet.

In dieser Arbeit wird der Wissenstransfer mittels Netzwerk behandelt, daher wird dieser im folgenden Abschnitt ausführlicher untersucht. Detailliertere Informationen zu den anderen Methoden und aktueller Literatur können im Überblick von Tan et al. (2018) aufgefunden werden.

5.5.1 Wissenstransfer mittels Netzwerk

Ein neuronales Netz zur Klassifikation kann in der Quelldomäne mit einer großen Datenmenge trainiert werden. Gut funktionierende Netzarchitekturen besitzen hierbei bereits eine hohe Klassifizierungsleistung. Es ist nun möglich, auch in der Zieldomäne vom generisch gelernten Wissen dieses trainierten Netzes zu profitieren. Hierfür können die Ausgaben von bestimmten Schichten des Netzes nach einer Dateneingabe als Merkmalsvektor genutzt werden (neuronales Netz mit Beschriftungen siehe Abbildung 5.3 a)). Die Ausgabe der letzten Zwischenschicht vor der Ausgabeschicht besitzt die höchste Datenabstraktion. Diese Merkmalsvektoren können für einen neuen Klassifikator genutzt werden, der andere Klassen besitzt.

Es kann außerdem das ursprüngliche neuronale Netz angepasst werden (siehe Abbildung 5.3). Hierzu wird die Ausgabeschicht mit einer neuen ersetzt, die die passende Anzahl Klassen (Neuronen) besitzt (siehe Abbildung 5.3 b)). Nur die Gewichtungen zwischen der letzten Zwischenschicht und der Ausgabeschicht müssen neu trainiert werden (siehe Abbildung 5.3 c)). Alle anderen Kantengewichte können beim Training „eingefroren“ (engl. frozen) werden. Optional ist es möglich, auch die restlichen Gewichtungen des ursprünglichen Netzes weiter anzupassen und somit auf das neue Problem zu optimieren (fine-tuning). Das Training wird in beiden Fällen stark beschleunigt, da generell wesentlich weniger Daten und auch weniger Trainingsepochen benötigt werden.

Diese Methode des Transferlernens bei tiefen neuronalen Netzen kann für verschiedene Arten von Daten verwendet werden. In dieser Arbeit wurde es für Audio- und Bilddaten untersucht, wobei die Audiodaten teilweise zunächst in Bilder (Spektrogramme) gewandelt wurden.

5.5.2 Untersuchung der Leistungsfähigkeit von Transferlernen

Einführung

Die folgenden beschriebenen Untersuchungen wurden in einem Konferenzpapier publiziert: (Seiderer, Dietz, Aslan & André, 2018). Alle Messwerte und Programme sind auf Github verfügbar ¹. Die Ergebnisse veranschaulichen, dass selbst auf mobiler Hardware und günstigen (< 60 €) Einplatinencomputern Transferlernen mit tiefen neuronalen Netzen mit praktikablem Zeitaufwand

¹<https://github.com/hcmlab/tf-transferlearning-benchmarks> (08.05.2019)

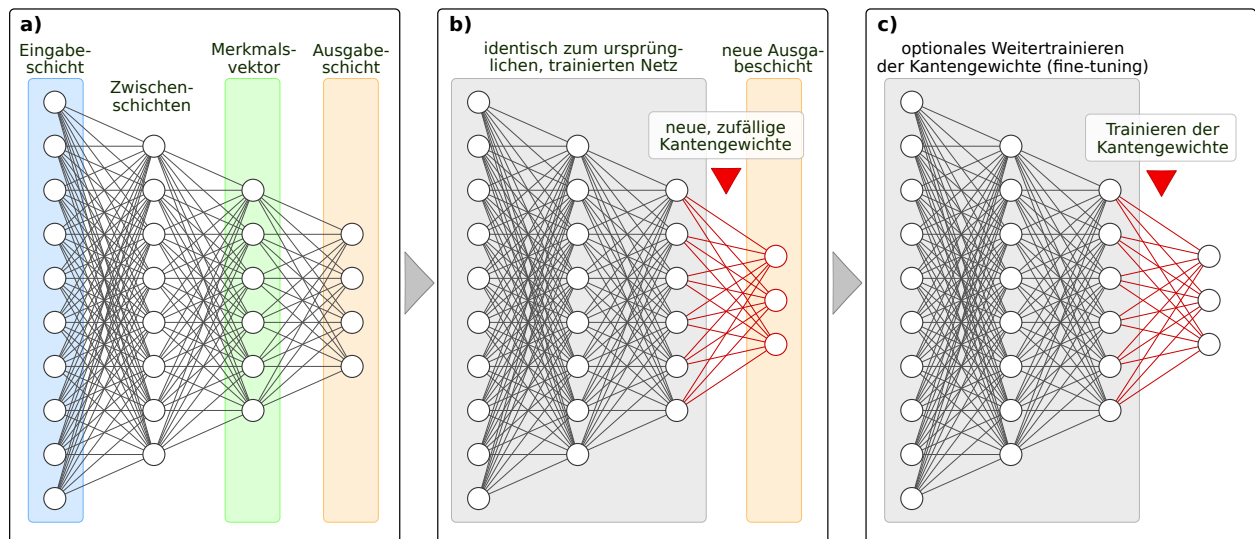


ABBILDUNG 5.3: Ablauf beim Transferlernen mittels Übertragen von Wissen mit neuronalem Netz.

und vergleichsweise wenigen Daten möglich ist. Hierzu wird die zuletzt behandelte Methode des Wissenstransfers mittels trainiertem neuronalem Netz genauer untersucht. Insbesondere beim Einsatz von mobiler Hardware muss eine Netzarchitektur gewählt werden, die einen akzeptablen Kompromiss zwischen Genauigkeit und Geschwindigkeit darstellt. Es wurde hier zudem untersucht, wie viele Bilder notwendig sind, um einen funktionierenden, neuen Klassifikator zu trainieren und inwiefern die Datenverteilung (gleichmäßig / ungleichmäßig) auf die verschiedenen Klassen eine Rolle spielt.

Kategorisierung

Für diese Untersuchung wird ein bereits trainiertes tiefes neuronales Netz wie in Abschnitt 5.5.1 angepasst. Ein auf 1000 Bildklassen trainiertes Netz soll für eine Bildklassifizierung von 5 anderen Klassen verwendet werden. Die Klassen und Datenverteilungen sind daher unterschiedlich. Die Merkmalsvektoren stammen in beiden Fällen aus dem unveränderten, vortrainierten Netz und sind daher in der Quell- und Zieldomäne identisch. Insgesamt sind die Domänen unterschiedlich: $\mathcal{D}_S \neq \mathcal{D}_T$, da zwar $\mathcal{X}_S = \mathcal{X}_T$ aber $P(X_S) \neq P(X_T)$. Es unterscheiden sich ebenfalls die Aufgaben: $\mathcal{T}_S \neq \mathcal{T}_T$, da $\mathcal{Y}_S \neq \mathcal{Y}_T$ und $f_S(\cdot) \neq f_T(\cdot)$. Beschriftungen der Daten sind in der Quell- und Zieldomäne vorhanden: \mathcal{D}_S und \mathcal{D}_T . Es handelt sich daher um überwacht (induktives) maschinelles Lernen; die Quell- und Zielaufgabe werden nicht gleichzeitig gelernt. Ein Teil des neuronalen Netzes aus der Quelle wird wiederverwendet, wodurch ein Wissenstransfer stattfindet.

Performanz mit verschiedenen mobilen / eingebetteten Geräten

Hardware Für die Untersuchung der Geschwindigkeit des Trainingsvorgangs wurden vier verschiedene Geräte mit ARM-Prozessor verwendet (siehe Tabelle 5.1). Zwei Smartphones und zwei eingebettete Systeme. Eine Version des System-on-Chip (SoC) vom Embedded System „Odroid HC1“ wurde für das Samsung Galaxy S5 (SM-G900H) genutzt, das ab April 2014 verfügbar war. Der Raspberry Pi 3 B ist ein sehr verbreiteter, kostengünstiger Einplatinenrechner, daher wurde auch dieser einbezogen. Die anderen Systeme erlauben einen Überblick über die Performanz, die low-end und mid-range Smartphones liefern können.

Software Die Betriebssysteme der Geräte unterscheiden sich, jedoch wurde die Testumgebung mittels „Chroot“ (Abkürzung für „change root“) vereinheitlicht, sodass auf allen Systemen ein

	Nexus 4	Nexus 6P	Raspberry Pi 3 B	Odroid HC1
SoC	Qualcomm Snapdragon S4 Pro	Qualcomm Snapdragon 810 v2.1	Broadcom BCM2837	Samsung Exynos5422
CPU	4x1,5 GHz Krait	4x2,00 GHz Cortex-A57 & 4x1,55 GHz Cortex-A53	4x1,2 GHz Cortex-A53	4x2,0 GHz Cortex-A15 & 4x1,4 GHz Cortex-A7
RAM	2 GiB, LPDDR2	3 GiB, LPDDR4	1 GiB, LPDDR2	2 GiB, LPDDR3
Speicher	NAND, 16 GiB	NAND, 64 GiB	SD card	SD card (OS) + SSD (data)
Host	Android 7.1 (LineageOS 14.1)	Android 8 (LineageOS 15.1)	Raspbian Stretch	Ubuntu Server 16.04.4
chroot	Linux Deploy: Debian Stretch armhf	Linux Deploy: Debian Stretch armhf	-	Docker: Debian Stretch (arm32v7/debian:stretch)
Jahr	11/2012	09/2015	02/2016	04/2014 (SoC)

TABELLE 5.1: Technische Daten der verwendeten Geräte mit ARM-Prozessor mit Angabe des Betriebssystems (host) und dem System der Testumgebung (chroot).

Debian-Stretch-armhf-Linux mit 32 bit genutzt werden konnte. Diese Umgebung erlaubt die Installation von Linux-Software, die für die ARM-Architektur geeignet ist. Dies ist auch unter Android möglich, da es sich hierbei um ein Linux-Derivat handelt. In dieser Arbeit wurde unter Android „Linux Deploy“² genutzt; es existieren jedoch auch andere Apps. „Docker“³ erlaubt unter Linux eine problemlose Nutzung einer Chroot-Umgebung in der Programme auch isoliert werden können. Die CPU-Architekturen müssen bei Chroot kompatibel sein, ansonsten wäre eine ineffiziente CPU-Virtualisierung notwendig.

Für die Tests wurde das weitverbreitete Framework „TensorFlow“⁴ von Google in Version 1.7 (nightly build⁵) mit der Python-API eingesetzt, das für den Raspberry Pi 3 B kompiliert wurde. Diese Version war auf allen Systemen mit allen Funktionen in einer Chroot-Umgebung lauffähig. Auf den Systemen wurde ausschließlich die CPU genutzt, da eine GPU- oder sonstige Hardware-Beschleunigung nicht verfügbar waren. Zur Ausführung der Tests wurde ein vorhandenes Skript aus dem TensorFlow-Tutorial⁶ entsprechend angepasst, sodass zusätzliche Informationen ausgegeben (z. B. Zeitstempel) und eine eigene Dateiliste angegeben werden konnten. Die Dateiliste legte fest, ob bestimmte Dateien zum Trainieren, Validieren oder Testen genutzt werden sollten. Die Dateiverteilung wäre ansonsten zufällig prozentual durchgeführt worden und hätte die Nachvollziehbarkeit erschwert.

Die durchgeführten Messungen ermöglichen eine Beurteilung der Größenordnung der benötigten Zeit beim Transferlernen auf dieser oder vergleichbarer Hardware. Optimierungen wären allerdings generell möglich gewesen.

Datensatz Als nachvollziehbaren Datensatz, auf den die Modelle angepasst werden sollten, wurde ein frei verfügbarer Datensatz mit verschiedenen Blumensorten genutzt. Dieser wurde bereits im erwähnten Tutorial eingesetzt. Es handelt sich hierbei um 3670 jpg-Bilder mit unterschiedlicher Qualität (mittlere Dateigröße: 62,0 kB), die von Flickr heruntergeladen wurden. Die Bilder sind auf fünf Klassen verteilt: Gänseblümchen (engl. daisy, 633 Dateien), Löwenzahn (engl. dandelion, 898 Dateien), Rosen (engl. roses, 641 Dateien), Sonnenblumen (engl. sunflowers, 699 Dateien) und Tulpen (engl. tulips, 799 Dateien).

Modelle Für die Experimente wurden verschiedene bereits trainierte Modelle genutzt (siehe Tabelle 5.2), die für die mobile Nutzung optimiert wurden (MobileNet v1 (Howard et al., 2017) / MobileNet v2 (Sandler et al., 2018)). Als Referenz wurde ein größeres Modell verwendet (Inception v3 (Szegedy et al., 2015)), das höhere Erkennungsraten liefern kann, jedoch wesentlich mehr

²<https://play.google.com/store/apps/details?id=ru.meefik.linuxdeploy> (16.05.2019)

³<https://www.docker.com/> (16.05.2019)

⁴<https://www.tensorflow.org/> (13.05.2019)

⁵git revision 5d33c1e49178aedbb459da7ce58eca710102c06b

⁶https://www.tensorflow.org/hub/tutorials/image_retraining (13.05.2019)

Rechenleistung benötigt. Die Modelle wurden mit dem Imagenet ILSVRC-2012-CLS-Datensatz trainiert, der aus 1,2 Millionen Bildern besteht, die über 1000 Klassen verteilt sind. Die Treffergenauigkeiten sind in Tabelle 5.2 angegeben. Die Werte stammen von mehreren Webseiten:⁷⁸⁹. Die Top-1-Treffergenauigkeit wurde nur über die Klasse, die vom Modell mit der höchsten Wahrscheinlichkeit ausgegeben wurde, bestimmt. Bei der Top-5-Treffergenauigkeit wurden die fünf Klassen mit den höchsten Wahrscheinlichkeiten herangezogen. Die letzte Zahl bei der Modellbezeichnung entspricht der Dimension der Eingabeschicht des Netzes. Beim Inception-v3-Modell sind es 299×299 Pixel. Bei den MobileNet-Modellen wird als vorletzter Wert die Größe des Netzes in Prozent angegeben. Die ausgewählten Netze konnten mittels „TensorFlow Hub“ direkt genutzt werden. Für die Tests wurden Netze gewählt, die vergleichsweise hohe Top-1- und Top-5-Treffergenauigkeiten aufwiesen. Das kleinste ausgewählte Modell war *MobileNet_v2_1.0_96*, das größte mobile Netz *MobileNet_v2_1.4_224*.

Model name	Top 1	Top 5
Inception_v3_299	78.0 %	93.9 %
MobileNet_v1_1.0_224	70.9 %	89.9 %
MobileNet_v1_1.0_128	65.2 %	85.8 %
MobileNet_v1_0.75_224	68.4 %	88.2 %
MobileNet_v1_0.75_128	62.1 %	83.9 %
MobileNet_v2_1.4_224	75.0 %	92.5 %
MobileNet_v2_1.0_224	71.8 %	91.0 %
MobileNet_v2_1.0_128	65.3 %	86.9 %
MobileNet_v2_1.0_96	60.3 %	83.2 %
MobileNet_v2_0.75_224	69.8 %	89.6 %
MobileNet_v2_0.75_128	63.2 %	85.3 %

TABELLE 5.2: Treffergenauigkeiten auf den ILSVRC-2012-CLS-Datensatz der Modelle, die für die Untersuchungen genutzt wurden.

Experiment: Geschwindigkeit und Ressourcen Das hierfür verwendete Programm führte folgende Schritte aus:

1. Suchen nach Bildern in einem angegebenen Ordner.
2. Herunterladen eines spezifizierten Modells von „TensorFlow Hub“.
3. Erzeugen und Speichern der Merkmalsvektoren.
4. Anpassung der Ausgabeschicht des Grundmodells für die neue Klassenanzahl.
5. Trainieren der Kantengewichte zwischen Merkmalsvektor und Ausgabeschicht und anschließendes Evaluieren.
6. Speichern des neuen Modells.

Es wurden für diesen Test die Standardparameter des ursprünglichen Skripts verwendet, sodass alle Bilder zufällig beim Training (80 % des Datensatzes), die Validierung (10 %) und Evaluation (10 %) eingesetzt wurden. Um Messfehler zu reduzieren, wurden alle Messungen dreimal durchgeführt und die Ergebnisse gemittelt. Alle Geräte waren mit Netzteilen mit konstanter Ausgangsspannung (Gleichspannung) verbunden. Bei den Android-Smartphones wurden somit Energiesparmaßnahmen verhindert, die die Leistungsfähigkeit der Geräte beeinflussen können. Die Geräte wurden bei normaler Zimmertemperatur (ca. 21 °C), ohne zusätzliche Kühlung, betrieben.

In Abbildung 5.4 sind die Ergebnisse des Laufzeittests für alle vier Geräte und alle 11 Modelle dargestellt. Die Farben geben an, wie viel Zeit eine bestimmte Phase des Skripts benötigte. Die

⁷<https://github.com/tensorflow/models/blob/master/research/slim/README.md> (06.04.2018)

⁸<https://github.com/tensorflow/models/tree/master/research/slim/nets/mobilenet> (06.04.2018)

⁹https://github.com/tensorflow/models/blob/master/research/slim/nets/mobilenet_v1.md (06.04.2018)

Laufzeit für die Erstellung der Merkmalsvektoren ist blau eingefärbt, die des Trainings orange und die der Evaluation grün.

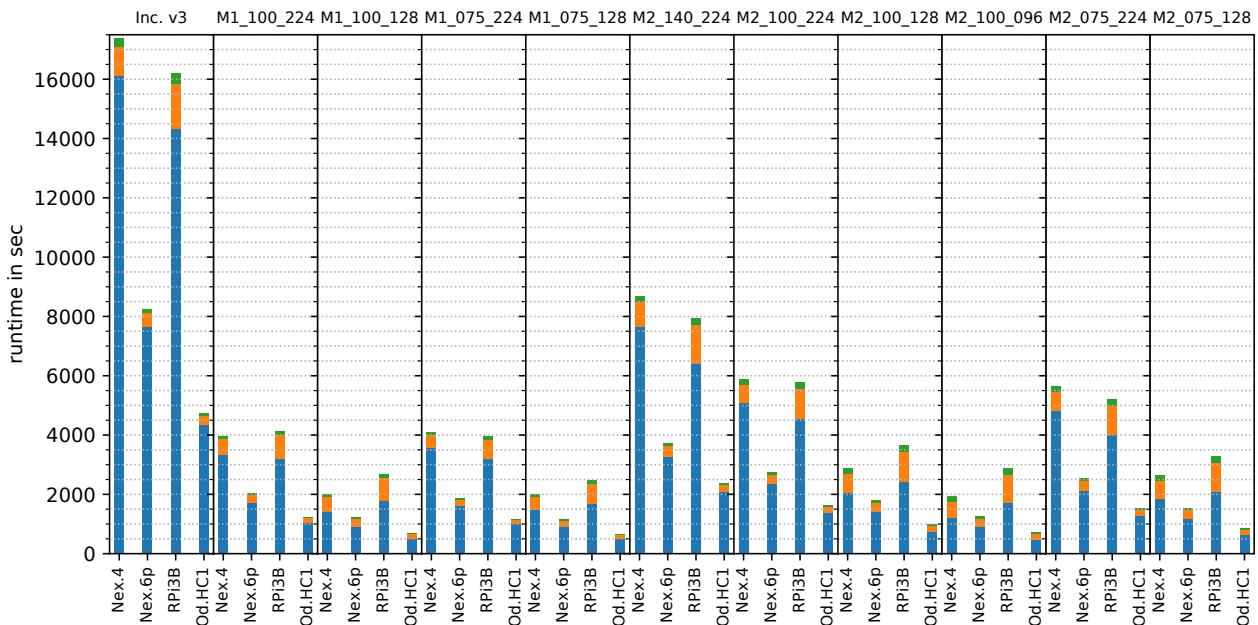


ABBILDUNG 5.4: Mittlere Laufzeit (drei Durchläufe) der Erzeugung der Merkmalsvektoren (blau), des Trainings (orange) und der Evaluation (grün) über alle Modelle mit dem Blumendatensatz. M1 = MobileNet v1, M2 = MobileNet v2, Inc v3 = Inception v3.

Generell benötigte die Erzeugung der Merkmalsvektoren die meiste Zeit, obwohl diese alle CPU-Kerne der Systeme verwendet. Beim Nexus-6P-Smartphone wurden nur die vier „großen“ Kerne des big.LITTLE-SoCs, beim Odroid-HC1-System wurden alle acht Kerne verwendet. Letzteres war im Vergleich immer das schnellste System. Für das Trainieren wurde allgemein wesentlich weniger Zeit benötigt, obwohl hierbei nur ein CPU-Kern genutzt werden konnte. Das Nexus-4-Smartphone und der Raspberry Pi 3 B benötigten insgesamt ähnlich viel Zeit. Obwohl der Raspberry eine aktuellere ARM-Architektur besaß, war bei diesem der Trainingsprozess langsamer als beim Nexus-4-Smartphone. Dies könnte am geringeren CPU-Takt des Raspberry Pi 3 B oder an der evtl. langsameren SD-Karte in Kombination mit weniger Arbeitsspeicher (Cache) gelegen haben.

Das komplexe Inception-v3-Modell benötigte bei allen Geräten, im Vergleich zu den anderen getesteten MobileNet-Modellen, mindestens zweimal so viel Zeit bei der Erzeugung der Merkmalsvektoren. *M1_100_128*, *M1_075_128* und *M2_100_096* konnten ähnlich schnell ausgeführt werden. Das schnellste Smartphone, das Nexus 6P, benötigte bei diesen Modellen ca. 15 min für die Erstellung der Merkmalsvektoren und 4 min 30 s für das Trainieren. Im Vergleich wurden beim Inception-v3-Modell 127 min (ca. 28,3-mal langsamer) für die Merkmalsvektoren und 8 min (ca. 1,8-mal langsamer) für das Training benötigt.

Die Dateigröße für alle Merkmalsvektoren-Dateien, die als unkomprimierte Text-Dateien gespeichert wurden, lag zwischen 77,3 MiB (Inception v3) und 21,7 MiB (*M1_075_128* / *M2_075_224*). Die ursprünglichen Bilddateien besaßen eine Größe von 222,1 MiB. Die gespeicherten neuen Modelle waren zwischen ca. 5,4 MiB (*M2_075_128* / *M2_075_224*) und 83,4 MiB (Inception v3) groß.

Es zeigt sich, dass bei dieser Dateianzahl weder die Laufzeiten noch die Dateigrößen für die mobilen / Embedded Systeme eine größere Herausforderung darstellten und somit der Einsatz ausgewählter tiefer neuronaler Netze mittels Transferlernen praktikabel ist.

Experimente: Bilderkennung mit zunehmender Bildanzahl Im Folgenden wurde untersucht, wie sich mittels Transferlernen trainierte Modelle mit verschiedenen Dateianzahlen, die gleichmäßig oder ungleichmäßig über Klassen verteilt sind, verhalten. Diese Untersuchungen erlauben

Rückschlüsse darüber, wie viele Daten pro Klasse nötig sind, um funktionierende Modelle zu erzeugen und inwiefern es bei kleinen Datenmengen notwendig ist, die Datenmenge für die Klassen auszugleichen.

Die drei Experimente wurden mit dem Modell *M2_100_128* durchgeführt, da dieses eine hohe Erkennungsrate auf dem ILSVRC-2012-CLS-Datensatz zeigte und im vorherigen Test verhältnismäßig geringe Laufzeiten auf den verschiedenen Geräten verursachte. Für diese Experimente wurde das veränderte Skript aus dem TensorFlow-Tutorial genutzt, das durch eigene Änderungen eine stabile Zuordnung zum Trainings-, Validierungs- und Test-Datensatz per Dateiliste erlaubte. Die Dateilisten wurden mit zusätzlichen eigenen Skripten generiert.

Bei den Experimenten wurden neu hinzugefügte Bilder zu 80 % in den Trainings- und 20 % in den Validierungs-Datensatz aufgenommen. 400 Bilder jeder Klasse, die nicht für das Training oder Validierung enthalten waren, wurden für den Test-Datensatz genutzt. Die Parameter des ursprünglichen Skripts wurden, bis auf die Trainingsepochenanzahl, die auf 1000 erhöht wurde, nicht variiert. Die Experimente wurden jeweils dreimal auf einem Windows-Desktop-PC mit TensorFlow 1.8 durchgeführt und die Ergebnisse gemittelt.

Experiment: Erhöhung der Bildanzahl einer Klasse Zunächst wurde eine Untersuchung mit unbalancierter Ergänzung von Daten durchgeführt. Es wurde hierbei simuliert, dass ein potenzieller Nutzer des Systems fünf neue Bilder einer neuen Klasse hinzufügt. In diesem Fall wurde die Klasse „Gänseblümchen“ ausgewählt. Es wurden immer fünf neue Bilder ergänzt, da dies eine realistische Bildanzahl ist, die bei einer Nutzung von einem Nutzer hinzugefügt werden könnten. Alle anderen Klassen beinhalteten konstant 150 Bilder. Die „Gänseblümchen“ Klasse wurde bis auf 150 Bilder erhöht. Nach jeder Datenergänzung wurde ein neues Modell mit Transferlernen trainiert, wobei immer das vortrainierte Modell *M2_100_128* die Basis bildete.

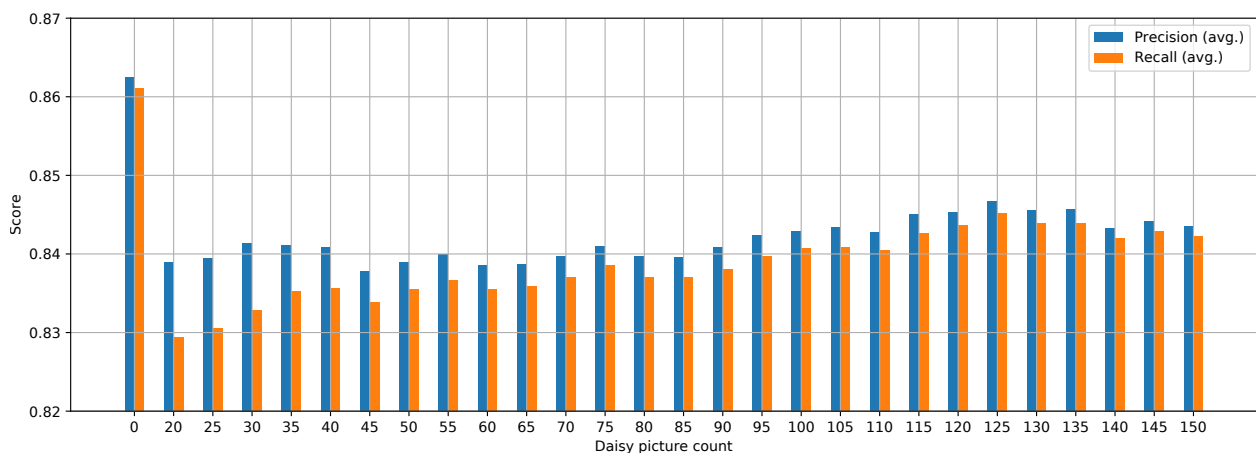


ABBILDUNG 5.5: Gemittelte Genauigkeiten (precision) und Trefferquoten (recall) der Modelle mit einer erhöhten Anzahl von Bildern mit Gänseblümchen (daisy) (80 % für Training, 20 % für Validierung, 400 Bilder pro Klasse für Test). Die anderen vier Bildklassen besaßen eine konstante Bildanzahl von 400 Bildern pro Klasse. Alle Trainings- und Evaluationsvorgänge wurden dreimal durchgeführt.

In Abbildung 5.5 sind die Resultate der Messungen dargestellt. Es zeigt die Genauigkeit und die Trefferquote für jede Modell-Evaluation mit 400 Bildern pro Klasse. Die Genauigkeiten und Trefferquoten wurden mit dem Macro-average-Mittelwert berechnet und sind über die drei Durchläufe gemittelt. Es ist erkennbar, dass nach dem Ergänzen der initialen 20 Bilder in die Klasse „Gänseblümchen“ die Genauigkeit und vor allem die Trefferquote von ca. 86 % auf ca. 84 % bzw. 83 % fielen, da eine neue Klasse dem Modell hinzugefügt wurde. Von diesem Punkt an erhöhte sich vor allem die Trefferquote mit zusätzlichen Bildern. Der Unterschied zwischen Genauigkeit und Trefferquote reduzierte sich, da sich die Bildanzahlen pro Klasse immer weiter annäherten.

Experiment: Erhöhung der Bildanzahl einer zufälligen Klasse Bei diesem Experiment wurden je fünf neue Bilder einer zufälligen Klasse hinzugefügt, wobei alle Klassen am Anfang keine Bilder enthielten. Damit ein Klassifizierungsmodell erstellt werden konnte, wurden zu Beginn mindestens zwei Klassen benötigt. Die Ergebnisse sind in Abbildung 5.6 dargestellt. Es ist erkennbar, dass am Anfang die Trefferquote und Genauigkeit instabil waren. Erst ab ca. der zehnten Iteration mit einer Gesamtbildanzahl von 50, beinhalteten alle Klassen Bilder. Mit erhöhter Anzahl an Bildern / Iterationen erhöhten sich die Genauigkeit und die Trefferquote, da die Unterschiede der Bildanzahlen zwischen den Klassen geringer wurden.

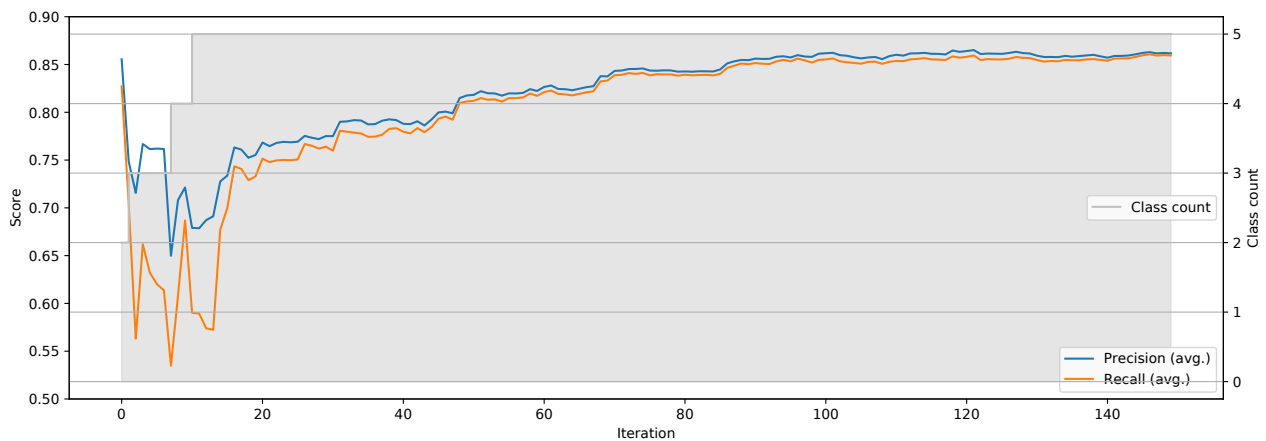


ABBILDUNG 5.6: Gemittelte Genauigkeiten (precision) und Trefferquoten (recall) der Modelle mit einer erhöhten Anzahl von Bildern (80 % für Training, 20 % für Validierung, 400 Bilder pro Klasse für Test). Neue Bilder wurden zufällig (stabiler „seed“ für den Zufallsgenerator) ergänzt, bis pro Klasse, 150 Bilder erreicht wurden. Die Anzahl der Klassen mit Bildern ist grau dargestellt. Alle Trainings- und Evaluationsvorgänge wurden dreimal durchgeführt.

Experiment: Erhöhung der Bildanzahl aller Klassen In diesem Fall wurde gleichmäßig die Bildanzahl pro Klasse um jeweils fünf Bilder pro Iteration erhöht. Die Evaluationsergebnisse sind in Abbildung 5.7 dargestellt. Der Unterschied zwischen Genauigkeit und Trefferquote ist hier gering. Beide Werte erhöhen sich bis ca. 60 Bilder pro Klasse.

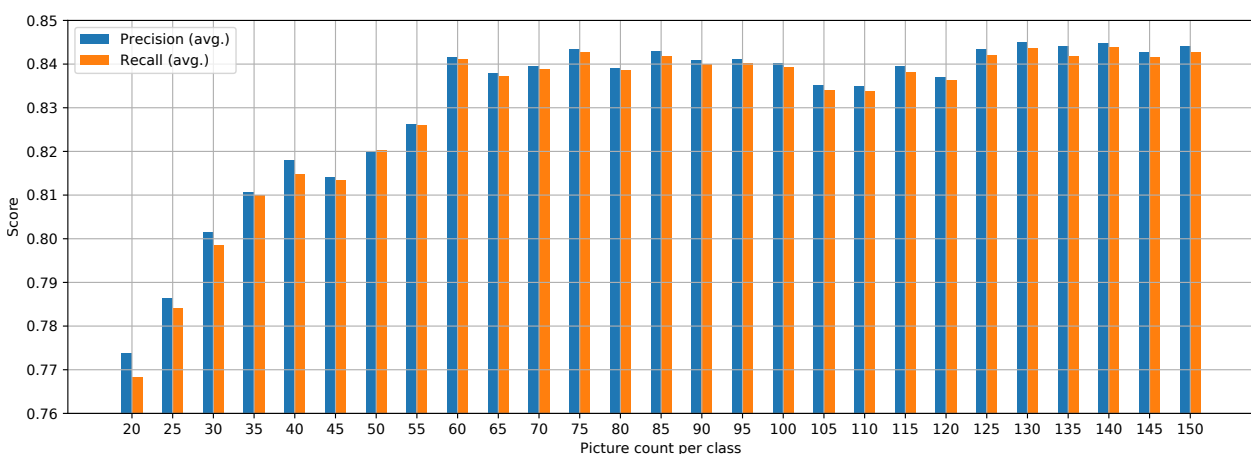


ABBILDUNG 5.7: Gemittelte Genauigkeiten (precision) und Trefferquoten (recall) der Modelle mit einer erhöhten Anzahl von Bildern pro Klasse (80 % für Training, 20 % für Validierung, 400 Bilder pro Klasse für Test). Alle Trainings- und Evaluationsvorgänge wurden dreimal durchgeführt.

Ergebnisse Die Resultate der Messungen zeigen, dass Transferlernen mit einem kleineren Datensatz und optimierten Modellen auf mobiler / Embedded Hardware in praktikablem Zeitrahmen genutzt werden kann. Neuere Hardware evtl. mit Beschleunigern sowie Software-Optimierungen können diesen Vorgang noch weiter beschleunigen. Größere Datensätze und Modelle bleiben jedoch leistungsfähigen Servern vorbehalten.

Die Tests mit dem inkrementellen Hinzufügen neuer Bilder veranschaulichen, dass bereits mit wenigen Daten relativ hohe Erkennungsraten erreicht werden können, selbst wenn die Klassen ungleichmäßig verteilt sind. Eine ungleichmäßige Verteilung würde natürlicherweise bei einer praktischen Nutzung auftreten. Es ist daher nicht zwingend notwendig, zum Trainieren ausschließlich Datensätze zu verwenden, bei denen pro Klasse ähnlich viele Bilder vorhanden sind. Falls durch zu große Unterschiede allerdings die Modellqualität stark reduziert würde, könnten verschiedene Verfahren zum Ausgleich der Bildanzahlen eingesetzt werden. Beispielsweise können vorhandene Bilder einer Klasse mit Bildeffekten verändert werden. Diese Funktionalität ist in vielen Frameworks, wie auch TensorFlow, bereits integriert.

Einer Nutzung von personalisierter Bilderkennung auf dem Endgerät des Nutzers spricht daher generell nichts entgegen. Es gibt jedoch einige Einschränkungen. Modelle sollten nur trainiert werden, wenn das Gerät an ein Netzteil angeschlossen ist und nicht verwendet wird, da ansonsten die Batterielaufzeit bei Smartphones dramatisch reduziert werden könnte. Die Bilderkennung mit neuen Modellen kann mit ungünstig aufgenommenen Bildern verschlechtert werden, daher muss ein praktikables System dies überprüfen und notfalls ungeeignete Änderungen verwerfen können. Zusätzlich kann diese Aufgabe, wie beim interaktiven maschinellen Lernen, mit einem entsprechenden Interface an den Nutzer weitergereicht werden. Falls das Ursprungsmodell jedoch generell nicht mehr geeignet ist, akzeptable Erkennungsraten mit dem Datensatz des Nutzers zu liefern, müsste dieses, soweit möglich, durch eine leistungsfähigere Netzarchitektur ersetzt werden.

5.5.3 Nutzung von Audiospektrogrammen

Insbesondere für die Bilderkennung existieren bereits Modelle, die auf sehr großen Datenmengen trainiert wurden und eine generische Repräsentation von Bilddaten gelernt haben. Diese umfassenden Modelle sind daher geeignet, um aus Bildern Merkmalsvektoren zu ermitteln. Auch für Audiodaten können sie genutzt werden, indem zunächst Spektrogramme erzeugt werden, womit das Audiosignal in eine visuelle Form umgewandelt wird. Bei Mel-Spektrogrammen wird im Vergleich zum Audiospektrogramm mittels der Mel-Skala eine Reduktion der Frequenzbänder durchgeführt, die dem subjektiven menschlichen Gehör nachempfunden ist (Stevens et al., 1937). In Abbildung 5.8 sind verschiedene Spektrogramme visualisiert, die z. B. direkt in das Modell *MobileNet_v2_1.4_224* eingespeist werden können, da sie eine Auflösung von 224×224 Pixeln besitzen. Bei den Spektrogrammen wirken sich die gewählte zeitliche Auflösung, Frequenzauflösung und auch die Farbpalette (Quantisierung) auf die Modellqualität aus. Beispielsweise konnten in der Arbeit von Amiriparian et al. (2017) Unterschiede bei den Farbquantisierungen festgestellt werden. Generell ist es möglich, mit den Merkmalsvektoren hohe Erkennungsraten bei Audiodaten zu erreichen, die häufig mindestens auf dem Niveau liegen, wie bei Modellen, die direkt auf den Audiodaten oder Audiomerkmale trainiert wurden.

In der vorliegenden Arbeit werden Audiodaten mittels Spektrogrammen in Abschnitt 7.3.2 klassifiziert. Ein weiteres Beispiel für die Anwendung bei einer Audioklassifikation zur Erkennung von Walen wurde in der eigenen Arbeit (Schiller, Huber, Lingenfelder, Dietz, [Seiderer](#) & André, 2019) präsentiert. Hierbei wurden zunächst Audiospektrogramme erzeugt und die Merkmalsvektoren mit drei aus der Bilderkennung bekannten tiefen neuronalen Netzen (Vggish, Inception v3, Resnet) berechnet. In der Arbeit konnte gezeigt werden, dass „explainable artificial intelligence“ oder auch XAI durch relevanzbasierten Maskierung dabei helfen kann, die Erkennungsraten noch zu optimieren.

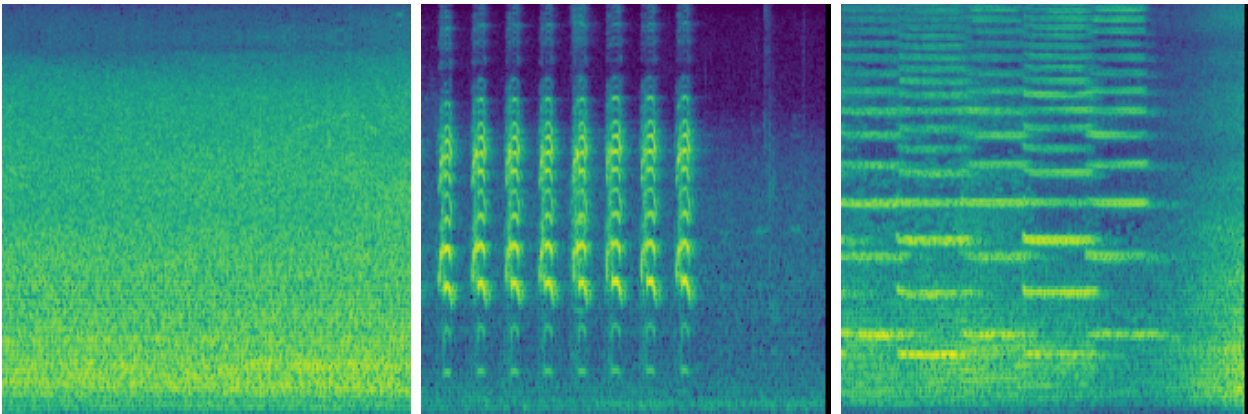


ABBILDUNG 5.8: Von links nach rechts: Mel-Spektrogramme von Straßenverkehr, Vogelgezwitscher und einer Sirene, die von der zeitlichen (224 Pixel) und Frequenzauflösung (224 Pixel) für das MobileNet_v2_1.4_224 optimiert sind. Es wurde die Farbpalette „viridis“ von Matplotlib genutzt.

5.5.4 Anwendungsbeispiele für Transferlernen

In dieser Arbeit wurde Transferlernen auch zur Verbesserung der Klassifikation von Lebensmitteln mit Daten von Gassensoren eingesetzt (siehe Abschnitt 8.3.3).

Weitere Beispiele für Transferlernen sind in der Literatur in den aktuelleren Übersichten von K. Weiss et al. (2016) und Tan et al. (2018) strukturiert aufzufinden. In der Mensch-Computer-Interaktion wird Transferlernen insbesondere bei Problemen, bei denen eine Personenabhängigkeit vorhanden ist, eingesetzt, um zu verhindern, dass ein neues Modell mit vielen Daten erstellt werden muss. Zu den besonders problematischen Biosignalen zählt z. B. EEG (Elektroenzephalografie) zur Ermittlung der Gehirnaktivität. Hierbei werden eine Vielzahl an Elektroden benötigt, deren Position pro Sitzung und Person variiert. Die Signale sind daher von der jeweiligen Sitzung und Person abhängig. In der Arbeit von Y.-P. Lin und Jung (2017) wurde daher Transferlernen genutzt, um den Aufwand einer individuellen Kalibrierung für ein System zur Emotionserkennung mittels EEG zu reduzieren. Ein vergleichbares Problem mit EMG-Signalen (Elektromyografie), die die Muskelaktivität widerspiegeln, wurde in der Arbeit von Prahm et al. (2019) behandelt. Bei der täglichen Nutzung von Prothesen für obere Extremitäten treten ebenfalls Verschiebungen der Elektroden (hier Thalmic Labs Myo-Armband – siehe auch Abschnitt 9.3.2) auf, deren Auswirkungen auf das Signal mithilfe von Transferlernen reduziert werden konnten. Auch die Emotionserkennung bei Sprache konnte mittels Transferlernen unter Verwendung verschiedener Korpora und Sprachen verbessert werden (Latif et al., 2018). In der Arbeit von Guo et al. (2018) wurde ein Modell zur Gesichtserkennung mittels Transferlernen für die Nutzung zur Erkennung von Lächeln angepasst, wodurch bessere Erkennungsraten erreicht werden konnten. Ein weiteres Beispiel für Transferlernen, in diesem Fall im Bereich „Human-Activity-Recognition“ mit Accelerometer-, Gyroskop- und Magnetometer-Sensordaten wurde im Artikel von Ding et al. (2019) präsentiert. Eine Personenunabhängigkeit des Klassifizierungsmodells (tiefes neuronales Netz) konnte durch unüberwachtes Transferlernen erreicht werden.

Es zeigt sich, dass Transferlernen für verschiedene Anwendungszwecke gut geeignet sein kann, um Probleme des maschinellen Lernens zu lösen und daher sehr aktiv erforscht wird. Allerdings existieren Limitationen, die bei der Anwendung berücksichtigt werden sollten. Beispielsweise kann Transferlernen, je nach verwendeten Datensätzen, Modelle verschlechtern, wenn diese zu unterschiedlich sind. Ein großes Problem ist zudem, dass ältere, bereits gelernte Aufgaben „vergessen“ werden können. Hierfür wurden in aktuellerer Literatur, wie beispielsweise von Kemker et al. (2018) bereits Messmethoden und Lösungsmöglichkeiten vorgestellt. Da Transferlernen auch Teil des Forschungsgebiets des lebenslangen Lernens ist, wurde dieses Problem auch im weiterführenden Literaturüberblick von Parisi et al. (2019) behandelt.

Teil III

Verfahren und Systeme zur Verbesserung des menschlichen Wohlbefindens

6	Emotionales, mentales und soziales Wohlbefinden	65
6.1	Erfassung von Stress, Lachen und sozialer Verbundenheit	66
6.2	Biofeedback mittels Tangible-Displays	69
6.3	Eigene Untersuchungen	70
6.4	Zusammenfassung des Kapitels	93
7	Umgebungsbezogenes Wohlbefinden	95
7.1	Erfassung von Umwelteinflüssen	97
7.2	Ambiente Systeme	103
7.3	Eigene Untersuchungen	105
7.4	Zusammenfassung des Kapitels	139
8	Physisches Wohlbefinden: Ernährung	141
8.1	Nahrungserfassung	143
8.2	Abfrage von Informationen zu Lebensmitteln	146
8.3	Eigene Untersuchungen	148
8.4	Zusammenfassung des Kapitels	198
9	Projekt CARE: Verbesserung des Wohlbefindens von Senioren	201
9.1	Konzept CARE-System	202
9.2	Empfehlungen zur Verbesserung des Wohlbefindens	207
9.3	Sensorik für das CARE-System	210
9.4	Entwicklung CARE-Prototyp 1	213
9.5	Entwicklung CARE-Prototyp 1.5 und 2	226
9.6	Entwicklung CARE-Prototyp 2 mit Gamifizierungen	236
9.7	Roboterassistent für Senioren	241
9.8	Zusammenfassung des Kapitels	247

6 Emotionales, mentales und soziales Wohlbefinden

Emotionales, mentales und soziales Wohlbefinden sind eng miteinander verbunden und werden im Rahmen dieser Arbeit folglich in diesem Kapitel gemeinsam betrachtet. Beispielsweise beeinflusst das Lachen in einer Gruppe, als positive Emotion, das soziale und emotionale Wohlbefinden. Dadurch, dass Lachen negative Gefühle abbauen kann, wird zudem das mentale Wohlbefinden gesteigert (siehe Abschnitt 6.1.2). Chronischer Stress hingegen kann dazu führen, dass negative Emotionen, wie z. B. Ärger, entstehen und reduziert hierdurch das mentale Wohlbefinden (siehe Abschnitt 6.1.1). In dieser Arbeit wurde als Indikator für positive Emotionen die Erkennung von Lachen und für die Abschätzung negativer Auswirkungen auf Emotionen Stresserkennung herangezogen. Zusätzlich wurde ein System entwickelt, das potenziell geeignet ist, Stress zu reduzieren und die soziale Verbundenheit zu erhöhen.

In Abbildung 6.1 ist ein vereinfachter Überblick über die in diesem Kapitel enthaltenen Untersuchungen und deren Bezug zum Wohlbefindensmodell dargestellt.

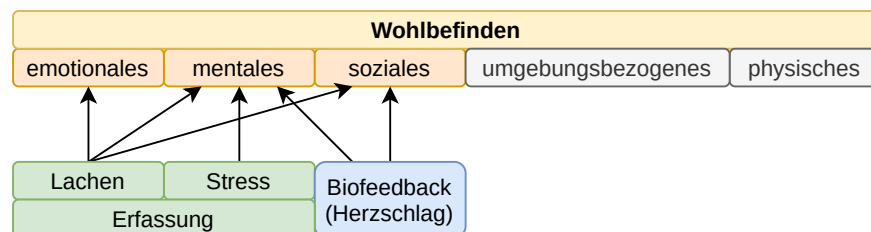


ABBILDUNG 6.1: Überblick über die durchgeführten Untersuchungen in diesem Kapitel.

Erfassung Die Erfassung von Informationen, die für das emotionale, mentale und soziale Wohlbefinden von Relevanz sind, stellt eine große Herausforderung dar. In dieser Arbeit wurden folgende Probleme behandelt, die bei der Erfassung auftreten können:

Problem 1: aufwändige Erfassung positiver und negativer Emotionen

Problem 2: Missachtung der Privatsphäre

Für eine möglichst zuverlässige Erkennung von Emotionen und Stress werden in der Literatur meist physiologische Daten, wie der Blutvolumenpuls oder der Hautleitwert herangezogen. Insbesondere der Hautleitwert kann allerdings nicht ohne direkten Hautkontakt erfasst werden. Zudem kann nicht immer davon ausgegangen werden, dass Fitnessarmbänder mit entsprechender Sensorik permanent getragen werden, da sie Personen eventuell stören oder das Anziehen vergessen wird. Daher müssen Lösungen entwickelt werden, um Emotionen und Stress möglichst unaufdringlich erfassen zu können (**Problem 1**). Eine in dieser Arbeit behandelte Lösungsmöglichkeit zur Erkennung positiver Emotionen besteht in der Nutzung kontaktloser mobiler und stationärer Sensoren. Beispielsweise kann Lachen per Audiosignal und Bewegungen erfasst werden. In Abschnitt 6.3.2 wird hierzu eine eigene Untersuchung bezüglich der mobilen Echtzeitklassifikation von Lachen mittels Smartphone vorgestellt. Lachen ist ein emotionales Ausdrucksverhalten. Je nach Auslöser

kann es ein Ausdruck verschiedener positiver Emotionen, wie beispielsweise Freude darstellen (Yim, 2016) und eine positive Gefühlsansteckung (engl. emotional contagion) (Hatfield et al., 1994) bei anderen bewirken. Hierdurch kann auch die soziale Verbundenheit verstärkt werden (Mauss et al., 2011). Es spielt daher hinsichtlich des emotionalen und sozialen Wohlbefindens eine durchaus beachtenswerte Rolle. Zusätzlich besitzt Lachen positive mentale Effekte (Yim, 2016) und somit einen Einfluss auf das mentale Wohlbefinden. Stress beeinflusst ebenfalls das mentale Wohlbefinden, da Stress durch mentale Überforderung ausgelöst werden kann. In der Untersuchung in Abschnitt 6.3.1 wurde die Erkennung von Stress im Büro behandelt. Als Besonderheit wurden hierbei, neben Referenzsensorik, stationär integrierbare Sensoren, wie drucksensitive Matten, genutzt.

Alle in diesem Abschnitt verwendeten Techniken und Systeme können den Schutz der Privatsphäre gewährleisten (**Problem 2**) (Abschnitt 1.2). So verwendete z. B. die Lacherkennung (Abschnitt 6.3.2) u. a. Audiodaten, die ausschließlich auf dem Smartphone in Echtzeit verarbeitet wurden. Bei der Stresserkennung (Abschnitt 6.3.1) konnte gezeigt werden, dass drucksensitive Matten das Potenzial besitzen, Biosignale zu ersetzen. Falls der Stress ausschließlich im Büro am eigenen Arbeitsplatz ermittelt werden soll, stellt dies eine Datenquelle mit wesentlich geringerem Datenschutzrisiko dar, als beispielsweise eine Videoüberwachung. Bei einem System für das Biofeedback des Herzschlags (Abschnitt 6.3.3) wurden quelloffene Software und ein eigener mobiler Hardware-Prototyp zur Datenerfassung und Verarbeitung entwickelt. Hierdurch mussten keine Biodaten an externe Anbieter übermittelt werden, wie dies aktuell bei vielen Fitness-Trackern der Fall wäre.

Biofeedback Beim Biofeedback handelt es sich um geschlossene Systeme (engl. closed-loop) bei denen Biosignale vom Nutzer erfasst und an diesen in einer bestimmten Form wieder zurückgegeben werden. Biofeedback kann beispielsweise zur Stressreduktion im Bereich der Meditation unterstützend eingesetzt werden (Shapiro et al., 2005), wodurch das mentale Wohlbefinden positiv beeinflusst wird. Zusätzlich kann eine Verbesserung des sozialen Wohlbefindens erreicht werden, wenn entsprechende Meditationsübungen in Gruppen durchgeführt werden, da hierdurch das Zusammengehörigkeitsgefühl / die soziale Verbundenheit gestärkt werden können.

Beim Biofeedback stellt die Auswahl einer geeigneten Ausgabemodalität ein wesentliches Problem dar. Dieses Problem wurde in Abschnitt 6.3.3 mit einem System untersucht, das ein eigenes entwickeltes Tangible-Display (künstliches Herz) als Ausgabegerät einsetzte, um den eigenen Herzschlag oder den des Partners fühlbar zu machen.

6.1 Erfassung von Stress, Lachen und sozialer Verbundenheit

6.1.1 Stress und dessen Auswirkungen auf das Wohlbefinden

Stress kann, wie Emotionen mittels Biosignalen von Computersystemen erfasst werden. Im Folgenden wird zunächst gemäß der Literaturzusammenfassung von Akbar et al. (2019) ein Überblick über Stress gegeben. In Stresssituationen werden die Ressourcen einer Person, wie z. B. Zeit, mentale oder finanzielle Ressourcen, überbeansprucht. Durch Stress werden mehrere physiologische Effekte ausgelöst, die vom Sympathikus und Parasympathikus stammen. Diese sind Teil des autonomen Nervensystems, das für die Steuerung der Körperfunktionen zuständig ist. Der Sympathikus steuert den Körper, um auf eine Gefahr oder Herausforderung zu reagieren, wodurch eine erhöhte Herzrate, Atemfrequenz und Muskelanspannung ausgelöst werden kann. Diese Reaktion ist auch bekannt als „Kampf-oder-Flucht-Reaktion“ (Kirschbaum, 2008). Während dieser werden alle Körperfunktionen, die nicht für das Überleben benötigt werden, unterdrückt. Der Parasympathikus ist der Gegenspieler, der den Körper in einen entspannten Zustand zurückführt. Wenn keine Stresssituation vorliegt, besteht ein Gleichgewicht zwischen beiden Systemen. Ein Ungleichgewicht kann über längere Zeiträume zu langfristigen Gesundheitsproblemen führen, kurzfristig können

negative Gefühle wie Ärger, Frustration, Angst, fehlende Motivation, Einschränkungen bei der Entscheidungsfindung und reduzierte Produktivität eine Folge sein. Während Stresssituationen wird das Stresshormon Cortisol ausgeschüttet, dessen Pegel einen zuverlässigen Wert für den Stresslevel darstellt. Der Cortisolspiegel kann allerdings derzeit nur aufwändig mittels Laboranalysen aus Blut- oder Speichelproben bestimmt werden und ist daher für eine Echtzeiterkennung bei Computersystemen nicht geeignet.

Generell sind für die Stresserkennung die Signalarten, die bei der Emotionserkennung eingesetzt werden, nutzbar. Die Übersichten von Greene et al. (2016) sowie von Akbar et al. (2019) besitzen einen Fokus auf die automatische Stresserkennung. Akbar et al. (2019) berücksichtigten allerdings primär Sensorik, die unaufdringlich ist. Der Überblick von Gradl et al. (2019) geht auf die Echtzeiterkennung mit den Daten von tragbaren Geräten ein, die derzeit verfügbar sind.

Für die Erkennung von Stress in der Büroumgebung sind beispielsweise die aktuellen Arbeiten von Koldijk et al. (2018), Yu et al. (2019) sowie Lawanot et al. (2019) relevant. In der Arbeit von Koldijk et al. (2018) wurden Computerinteraktionen geloggt, der Gesichtsausdruck per Kamerabild, die Körperhaltung mittels Kinect und physiologische Sensordaten (Herzrate, Hautleitwert) von spezieller Hardware erfasst. Die Daten wurden anschließend mit klassischen Verfahren des maschinellen Lernens aus dem WEKA-Toolkit analysiert und Stress konnte mit dem Datensatz zuverlässig erkannt werden. In Lawanot et al. (2019) wurden für die Stresserkennung ausschließlich Kamerabilder und tiefe neuronale Netze eingesetzt. Bei diesen beiden Arbeiten wurden Kamerabilder genutzt, die insbesondere an Arbeitsplätzen in einigen Ländern, wie Deutschland, nicht generell zulässig wären. Physiologische Daten sind allerdings für eine Erkennung ausreichend und wurden auch in (Koldijk et al., 2018) und vielen anderen Arbeiten für Emotions- und Stresserkennung erfolgreich genutzt. Eine Privatsphäre schützende Methode ist durch die Verwendung von drucksensitiven Matten möglich. Ein aktuelles Beispiel ist hierbei die Arbeit von Yu et al. (2019), bei der Büroaktivitäten und die Herzrate aus den Signalen einer drucksensitiven Matte (1x1 Matrix) ermittelt wurden, die für eine Abschätzung des Stresslevels grundsätzlich nutzbar wären. Um die Daten nicht nur am Arbeitsplatz messen zu können, wäre zudem eine Inklusion tragbarer Sensorik notwendig, deren Nutzung beispielsweise in der Arbeit von Montesinos et al. (2019) untersucht wurde.

6.1.2 Lachen als positiver Emotionsausdruck und dessen Auswirkungen auf das Wohlbefinden

Die Erfassung und der Ausdruck von Emotionen bei Computern zählt zum Forschungsgebiet „Affective Computing“ (Picard, 2000). Für die generelle Erkennung von Emotionen wurden in der Literatur verschiedene Biosignale genutzt. Hierzu zählen die Sprache (Audiosignal) (Xie et al., 2019), Gehirnaktivität (EEG-Signal) (Ozdemir, Degirmenci et al., 2019), Gesichtsausdruck (Video-Signal) (Ozdemir, Elagoz et al., 2019) bzw. Gesichtsmuskelaktivität (EMG-Signal) (Mithbavkar & Shah, 2019) sowie Herzschlag (BVP-Signal) und Hautaktivität (EDA-Signal¹) (Domínguez-Jiménez et al., 2020). Ein umfangreicherer Überblick wird in der Arbeit von Shu et al. (2018) gegeben. Einige Lösungen nutzen multimodale Datenquellen, wie z. B. in der Arbeit von Wagner, Lingenfelder und André (2015), bei der Audio und Video eingesetzt wurden. In diesem Fall ist die Modalität „Sprache“ nicht dauerhaft verfügbar, diese kann allerdings unterstützend genutzt werden. Bei einer mobilen Erfassung sind in den meisten Fällen einige der Modalitäten nicht permanent verfügbar. Eine Nutzung der Gehirnaktivität ist derzeit beispielsweise nur sehr eingeschränkt mobil möglich. Sprache und der Gesichtsausdruck können nur sporadisch erfasst werden. Als dauerhaft verfügbare Basis einer Erkennung sind daher Informationen über den Herzschlag oder die Hautaktivität gut geeignet, falls ein entsprechendes Armband getragen wird. Um die Erfassung zu verbessern, können weitere Modalitäten hinzugenommen werden.

¹Elektrodermale Aktivität (engl. electrodermal activity); häufig auch: „galvanic skin response“ (GSR)

Laut Cosentino et al. (2016) ist Lachen als nonverbales, soziales Verhalten für verschiedene Disziplinen von Relevanz, da es aus unterschiedlichen Gründen ausgelöst werden kann: als Antwort auf externe physische, kognitive oder emotionale Stimulation, um soziale Interaktionen auszuhandeln und pathologisch, als Folge einer neurologischen Schädigung. Lachen als Ausdruck positiver Emotionen besitzt vorteilhafte physiologische und mentale Auswirkungen, daher spielt es auch beim Wohlbefinden eine wesentliche Rolle. Gemäß der Arbeit von Savage et al. (2017) können durch Lachen Stresshormone abgebaut sowie das Immunsystem und die Gesundheit gestärkt werden. Zudem konnte in Lernumgebungen durch Humor die soziale Bindung zwischen Schülern und Lehrern verbessert werden, sodass hier auch positive Effekte auf das soziale Wohlbefinden in Gruppen auftreten können. Das mentale Wohlbefinden kann durch Lachen verbessert werden, da hierdurch negative Gefühle reduziert werden können (Yim, 2016).

Gemäß dem Literaturüberblick von Cosentino et al. (2016) treten beim Lachen drei Hauptmerkmale auf: 1) Lachgeräusch, 2) Gesichtsausdruck, 3) Bewegung des gesamten Körpers. Die Autoren erwähnen zudem verschiedene Veröffentlichungen, die Lachmerkmale, teilweise in Kombination, für die Klassifizierung einsetzen.

6.1.3 Soziale Verbundenheit und deren Auswirkungen auf das Wohlbefinden

Soziale Verbundenheit wird in der Arbeit von Haslam et al. (2015) als das Gefühl der Zugehörigkeit und Bindung, die Menschen in Bezug auf Einzelpersonen und Personengruppen fühlen definiert. Sie stellt einen bedeutsamen Indikator für Wohlbefinden und Gesundheit dar. Im Überblick von Haslam et al. (2015) wird eine Vielzahl an Publikationen zusammengefasst, bei denen positive Auswirkungen auf die Lebenserwartung, mentale und kognitive Gesundheit festgestellt wurden. Gemäß Haslam et al. (2015) spielen hierbei psychologische und physiologische Mechanismen eine Rolle. Beispielsweise kann soziale Unterstützung dazu führen, dass Stress besser abgebaut werden kann, der schädliche psychologische und physiologische Effekte besitzen kann. Laut den Autoren spiegelt die Anzahl der sozialen Kontakte allerdings nicht deren Wirkung wider, sondern es spielt auch der Umfang der sozialen Interaktion eine Rolle. Viele der im Überblick enthaltenen Arbeiten beziehen sich auf Senioren, die von sozialer Isolation häufiger betroffen sind, da sie in vielen Fällen alleinstehend sind und infolge körperlicher Einschränkungen oder Krankheit kaum Kontakte zu Mitmenschen haben. Zusätzlich können sie häufig moderne digitale Kommunikationsmöglichkeiten nur eingeschränkt nutzen. Eine Verbesserung des sozialen Wohlbefindens von Senioren sollte daher auch beim CARE-Projekt (siehe Kapitel 9) mit IoT-Technologien ermöglicht werden.

6.1.4 Eigene Arbeiten

Zur **Erkennung von Lachen** wurde in der eigenen Arbeit in Abschnitt 6.3.2 eine multimodale Echtzeitklassifizierung entwickelt. Es wurden Audio- und Bewegungsdaten genutzt, die ein Smartphone bereitstellte, das in einer Brusttasche platziert wurde. Die eigene Arbeit zeichnet sich gegenüber bisherigen Veröffentlichungen dadurch aus, dass die Echtzeitklassifizierung auf demselben Gerät (Smartphone von 2013) erfolgte. Das System zeigte, dass die Hardware eines Smartphones hierfür ausreichend Leistung liefern kann und mehrere Modalitäten die Erkennung von Lachen verbessern können. Insbesondere falls eine der Modalitäten ausfällt, wie z. B. durch laute Umgebungsgeräusche oder Bewegung, kann eine grundsätzliche Erkennung dennoch durchgeführt werden. Mittlerweile können allerdings auf aktuellen Smartphones auch deutlich komplexere Klassifizierungen durchgeführt werden, wobei zudem längere Akkulaufzeiten möglich sind. Die verfügbare Sensorik hat sich allerdings nur geringfügig verändert, da diese durch die Platzierung der Smartphones am Körper beschränkt wird. In der Literatur sind für die multimodale Echtzeiterkennung von Lachen mit Audio- und Bewegungsdaten auf Smartphones jedoch keine neueren Arbeiten aufzufinden. Eine Smartwatch-Anwendung zur Echtzeit-Lacherkennung, die

tiefe neuronale Netze verwendet wird in der neueren Arbeit von Hagerer et al. (2017) präsentiert. Diese nutzt jedoch ausschließlich Audiodaten.

Zur **Erfassung des Stresslevels** wird in Abschnitt 6.3.1 die Nutzung von kommerziellen, preisgünstigen, drucksensitiven, resistiven Boden- und Stuhlmatten untersucht. Neben den klassischen Biosignalen bei der Stresserkennung (Blutvolumenpuls und Hautleitwert) konnten die Daten der Matten, trotz ihrer vergleichsweise geringen temporären und spatialen Auflösungen, erfolgreich für eine Klassifizierung eingesetzt werden. Der Vorteil gegenüber Biosignalen liegt darin, dass die Druckdaten z. B. am Arbeitsplatz aufgezeichnet werden können, ohne externe Sensorik einsetzen zu müssen. Es wird zudem ein möglicher Missbrauch der Daten stärker eingeschränkt, als wenn für diese Aufgabe beispielsweise Videodaten eingesetzt würden. Eine Alternative könnte ein permanent getragenes Fitnessarmband mit Sensoren für Blutvolumenpuls und Hautleitwert darstellen, wenn die Daten ausschließlich auf dem Gerät des Nutzers verarbeitet werden. In aktueller Literatur sind keine vergleichbaren Studien mit kostengünstigen Druckmatten zur Stresserkennung aufzufinden, bei denen ein Vergleich mit anderen Modalitäten unter Nutzung identischer Stress-Stimuli durchgeführt wurde.

In Abschnitt 6.3.3 wurden spezielle mobile Hardware- / Software-Prototypen für **Biofeedback** mit der Bezeichnung „PiHeart“ entwickelt, die u. a. einen modernen, optischen Sensor zur Messung des Blutvolumenpulses enthielten. Dieser Sensor konnte an einem Finger platziert werden. Eine eigene Software wurde entwickelt, um aus dem BVP-Signal die Herzrate im Frequenzbereich in Echtzeit zu bestimmen. Alternativ wäre auch eine Kombination aus kommerziellem Smartphone und Fitness-Tracker denkbar gewesen. Allerdings sind die hierbei genutzten Bluetooth-Verbindungen häufig nicht stabil (der eigene Körper kann die Funkverbindung unterbrechen) und es existieren nur wenige Fitness-Tracker, die mit eigener Software ein dauerhaftes Auslesen der Herzrate ermöglichen. Da für die tangible Ausgabe ein eigens angefertigter Aktuator-Prototyp benötigt wurde, bot es sich an, ein kompaktes, mobiles Gesamtsystem aus Aktuatorik, Sensorik, Einplatinencomputer und Stromversorgung zu entwickeln. Dieses konnte mit geringem Transport- und Installationsaufwand für eine Feldstudie eingesetzt werden und wäre zudem für weitere Anwendungen leicht modifizierbar gewesen. Verwandte Arbeiten werden in Abschnitt 6.3.3 diskutiert. Direkt vergleichbare Prototypen, deren Auswirkungen auf das Nutzererlebnis in einer umfangreichen Studie untersucht wurden, sind in aktueller Literatur jedoch nicht aufzufinden.

6.2 Biofeedback mittels Tangible-Displays

Im Literaturüberblick von Yu, Funk et al. (2018) ist eine Vielzahl an Arbeiten bezüglich der Reduktion von Stress mittels Biofeedback enthalten. Die meisten Studien in diesem Überblick nutzten Herzsignale für Systeme zur Reduktion von Stress. Der Großteil der Biofeedback-Systeme setzten grafische und numerische Ausgaben für das Feedback ein. Als Ausgabegerät kamen hierbei zu meist PCs mit Bildschirm zum Einsatz. Tangible-Displays spielten kaum eine Rolle, was auch daran liegen dürfte, dass diese nur mittels spezieller Hardware erstellt werden können und entsprechend mit höherem finanziellen Aufwand erst konstruiert werden müssen. Moderne, automatisierte Werkzeuge wie 3D-Drucker oder CNC-Maschinen können bei der Entwicklung von Prototypen jedoch sehr behilflich sein und erlauben die Erstellung komplexerer greifbarer Konstruktionen.

Ein Biofeedback-System (PiHearts) bei dem die Systemausgabe auf einem neuartigen Tangible-Display basierte, wird in Abschnitt 6.3.3 vorgestellt. Hierbei wurde ein u. a. vom Autor konstruiertes künstliches Herz zur Datenausgabe verwendet, das beim Halten in der Hand fühlbare Herzschläge simulierte. Es konnte der Herzschlag derselben Person sowie der einer anderen ausgegeben werden. Das künstliche Herz sollte hinsichtlich Aussehen, Gewicht, Struktur und Bewegungen einem biologischen, menschlichen Herzen ähneln. Allerdings sollte vermieden werden, dass Anwender ein zu realitätsnahes, künstliches Herz ablehnen. Daher wurde beispielsweise auf eine schleimige, feuchte Hülle und realistische Farben verzichtet. Der Einfluss des Prototyps auf

die Nutzererfahrung beim Ansehen von Filmen wurde in einer Feldstudie mit 60 Teilnehmern untersucht. Hierbei konnte festgestellt werden, dass ein tieferes Eintauchen in einen Film mithilfe des Biofeedback-Systems möglich ist. Zudem konnte eine stärkere soziale Verbundenheit mit einem Partner auftreten, wenn die Herzdaten mit dem jeweiligen Partner geteilt wurden.

6.3 Eigene Untersuchungen

6.3.1 Stresserkennung bei der Arbeit

Einleitung

Insbesondere in der Arbeitsumgebung kann Stress auftreten, der dort vor allem entsteht, wenn die Anforderungen an eine Arbeitskraft zu hoch werden (Demerouti et al., 2001). Falls sich Personen nicht ausreichend erholen können, kann chronischer Stress mentale Krankheiten auslösen. Daher kann es für Büroarbeiter hilfreich sein, von einem System automatisch auf einen längerfristig erhöhten Stresslevel hingewiesen zu werden, um rechtzeitig Pausen einzulegen. In einer Arbeitsumgebung sollte jedoch ein entsprechendes System nicht ausschließlich zwischen einem entspannten und gestressten Zustand unterscheiden, da auch der erwünschte konzentrierte Geisteszustand auftritt und nicht differenziert berücksichtigt werden könnte.

In dieser Arbeit wurde untersucht, inwiefern der Stresslevel von Büroarbeitern erfasst werden kann, ohne eine Audio- oder Videoüberwachung zu benötigen, die besonders schwerwiegend in die Privatsphäre eingreifen würden. Idealerweise sollten Sensoren genutzt werden, die in die Büroumgebung integriert werden können. Für diesen Zweck wurden drucksensitive Matten auf dem Bürostuhl sowie dem Boden untersucht. Als Referenz wurden physiologische Signale (BVP, EDA) und die Augenbewegungen aufgezeichnet. Stress und Emotionen besitzen direkten Einfluss auf das autonome Nervensystem, dessen Aktivitäten zum Großteil unwillkürlich sind. Physiologische Daten, die ein Resultat der Aktivitäten des autonomen Nervensystems darstellen, werden daher als zuverlässige Signale bei der Stress- und Emotionserkennung angesehen (Lang et al., 1998).

Es wurde im Rahmen dieser Arbeit eine Laborstudie durchgeführt, um kontrolliert Daten aufzuzeichnen, die auch den Unterschied zwischen einem gestressten und geistig konzentrierten Zustand von Personen berücksichtigten. Die Datenauswertung erfolgte anschließend mittels maschinellem Lernen und erlaubte es, zu erkennen, welche Kombination von Modalitäten für jeden Stresslevel die besten Resultate lieferte.

Die Untersuchung wurde in der eigenen Arbeit (Nakashima, Kim, Flutura, [Seiderer](#) & André, 2016) publiziert.

Verwandte Arbeiten

Stress-Stimuli Für die Aufzeichnung von annotierten Daten war es notwendig, gezielt Stress auszulösen. Hierfür kann der „Stroop-Test“ eingesetzt werden, der beispielsweise in den Arbeiten von Akbar et al. (2019), Calibo et al. (2013) und K. Frank et al. (2013) genutzt wurde. Es existieren beim Stroop-Test zwei verschiedene Versionen: der kongruente und inkongruente. Beim kongruenten Test wird z. B. der Farbname „gelb“ in Textform angegeben und dieser Text besitzt auch die entsprechende Farbe. Beim Inkongruenten wäre die Farbe desselben Texts z. B. rot. Der Studienteilnehmer muss in beiden Fällen mit der Farbe des Texts antworten. Eine Inkonsistenz zwischen geschriebenem Wort und dessen Farbe führt hierbei zu Stress. Diese Art von Test besitzt den Vorteil, dass der erzeugte Stress bei allen Teilnehmern vergleichbar hoch ist. Dies ist beispielsweise bei mathematischen Berechnungen nicht gegeben, da hier die individuellen Fähigkeiten eine Rolle spielen.

In einigen Studien wurden auch realistischere Aufgaben zur Erzeugung von mentalem Stress eingesetzt, die mit typischer Büroarbeit vergleichbar sind. Hierzu wurden Text-Transkriptionen,

das Erfassen von Informationen und ausdrucksvolles Schreiben als mentale Stressoren entwickelt (Hernandez et al., 2014). Die Stressoren wurden so gestaltet, dass sie typisch für Büroarbeiten sind. Hierzu zählen Zeitdruck und Unterbrechungen (durch Anrufe oder E-Mails). Auch in den Untersuchungen von Hart und Staveland (1988) für die Entwicklung des NASA-Task-Load-Index-Fragebogens (NASA-TLX) spielte hoher Zeitdruck eine bedeutsame Rolle bei der Bestimmung der subjektiven Arbeitslast (engl. workload).

Nach der Erzeugung von Stress wird dieser, wie beispielsweise in (Hernandez et al., 2014), durch erholsame Videos abgebaut, um daraufhin den Grundzustand (entspannt) wiederherzustellen.

Modalitäten In den Studien von Arnrich et al. (2009) und Meyer et al. (2010) konnte festgestellt werden, dass Drucksensorik nutzbar ist, um Stress mithilfe der Körperhaltung zu erkennen. Zunächst wird hierfür die Druckverteilung auf dem Stuhl einer Person gemessen und hieraus die Körperhaltung berechnet. Diese kann für die Bestimmung der Emotion und von Stress herangezogen werden. In der Veröffentlichung von Arnrich et al. (2009) konnte Stress direkt aus der Druckverteilung erkannt werden. Die Arbeit von Giraud et al. (2013) zeigt, dass es auch möglich ist, die Druckverteilung am Boden für eine Emotionserkennung zu nutzen.

Physiologische Signale haben in vielen Studien wie z. B. von Akbar et al. (2019), Plarre et al. (2011), Renaud und Blondin (1997) und Zhai und Barreto (2006) bewiesen, dass sie für die Erkennung von Emotionen oder Stress geeignet sind. Der Hautleitwert ist hierfür eines der am besten geeigneten Signale. Das BVP-Signal und die hieraus abgeleitete Herzfrequenz und Herzratenvariabilität werden ebenso häufig eingesetzt. Die unwillkürlichen Augenbewegungen, wie Sakkaden oder Fixierungen reflektieren die Aktivitäten des zentralen Nervensystems und können daher ebenfalls bei der Emotions- und Stresserkennung hilfreiche Informationen liefern (Di Stasi et al., 2013).

Studie

Stress-Stimuli Für die Studie wurden als Stressoren der Stroop-Test (siehe Abbildung 6.2) sowie die Informationsrecherche (siehe Abbildung 6.3) ausgewählt. Der kongruente Stroop-Test diente zur Herbeiführung eines konzentrierten Zustands, der Inkongruente sollte Stress verursachen. Bei der Informationsrecherche mussten 10 Fragen mittels eines Texts mit ca. 900 Wörtern beantwortet werden. Ohne Zeitlimitierung sollte diese realitätsnähere Aufgabe einen konzentrierten Zustand hervorrufen, unter Zeitdruck (3 Minuten) sollte Stress ausgelöst werden.



ABBILDUNG 6.2: Stroop-Test. Links kongruent (korrekte Antwort: grün), rechts inkongruent (korrekte Antwort: gelb).

Datenaufzeichnung Der Studienaufbau ist in Abbildung 6.4 dargestellt. Es wurden drei Druckmatten von SensingTex² zur Erfassung von Körperbewegungen verwendet. Eine wurde an der

²<http://sensingtex.com/> (07.03.2020)

Bitte lesen Sie den Text gründlich und beantworten Sie dann die Fragen darunter

Verbleibende Minuten: 1:25

Mandarine bezeichnet sowohl eine [Zitruspflanze](#) als der [Familie](#) der [Rautengewächse](#) als auch die orangefarbene [Frucht](#) derselben. Die Frucht verdankt ihren Namen wahrscheinlich den [Mandarinern](#), hohen chinesischen Staatsbeamten.^[1]

Herkunft

Die Herkunft der Mandarinen wird im Nordosten Indiens oder Südwesten Chinas vermutet. Die Pflanzen werden in China schon seit einigen tausend Jahren kultiviert, die erste sichere Erwähnung stammt aus dem 12. Jahrhundert v. Chr. Von der Ursprungsregion breitete sich die Mandarine über Südostasien und Indien aus. Um das 1. Jahrtausend n. Chr. wurde die Mandarine schon in vielen südlichen Präfekturen Japans kultiviert. Die ersten Mandarinen, die nach Europa eingeführt wurden, kamen 1805 mit Sir [Abraham Hume](#) aus [Kanton](#) China nach England („Kantoner Apfelsine“).^[2] Aus einer dieser ersten zwei Sorten hat sich später die „Mittelmeer“-Mandarine entwickelt.

Mandarinenähnliche

[Clementinen](#) und [Satsumas](#) sind kernlose Zitrusfrüchte, die als [Hybriden](#) zwischen [Orange](#) und Mandarinen entstanden sind.^[3] Sie werden sowohl botanisch als auch warenkundlich von der Mandarine unterschieden. Außerdem gibt es [Minneola](#), eine Kreuzung aus Mandarine und Grapefruit.^[4] Die Orange ist selbst aus einer Kreuzung der Mandarine hervorgegangen, und zwar mit der [Pampelmuse](#).^[5]

Beschreibung

Mandarinen sind die variabelste und größte Gruppe der Zitruspflanzen was Fruchtform, Größe, Geschmack der Frucht und Habitus der Pflanzen betrifft. Sie sind in der Regel kleine, immergrüne [Bäume](#). Die Zweige sind nur mit wenigen, kleinen Dornen besetzt. Die Blätter sind lanzettlich, auf beiden Seiten spitz zulaufend. Der Blattstiel ist nur undeutlich von der Blattspreite abgesetzt, die Flügel am Blattstiel sind nur als schmale Linie erkennbar. Die Blattränder sind undeutlich gekerbt. Die [Blüten](#) stehen einzeln oder in wenigblütigen Blütenständen in den Blattachseln. Die [Kelchblätter](#) sind verwachsen, die fünf weißen [Blütenblätter](#) frei. Die 20 bis 25 [Staubblätter](#) sind in mehreren Gruppen miteinander verwachsen. Der [Griffel](#) ist lang und schmal. Die Früchte ([Hesperiden](#)) der Mandarine sind wesentlich kleiner als [Orangen](#), sie schmecken weniger sauer als die Orange und weisen ein unverwechselbares intensives, komplexes Aroma auf. Ihre Haut lässt sich im Vergleich zu anderen Zitrusfrüchten leichter abschälen, und sie lässt sich auch besonders leicht in von außen trockene Segmente zerteilen, wodurch man sie gut mit den Fingern schälen und essen kann. Jede Frucht besteht aus etwa neun Segmenten, die mit orangefarbenen [Saftschläuchen](#) gefüllt sind. Jedes Segment ist von einem dünnen Häutchen ([Endokarp](#)) umgeben, die ganze Frucht von einer zweigeteilten Schale. Die innere Schicht der Schale ist weiß ([Mesokarp](#), Albedo), die äußere grün bis orange ([Exokarp](#), Flavedo). Die weiße Schicht ist bei der Reife auf ein Netz aus Fasern reduziert, darauf bezieht sich der wissenschaftliche Name "reticulata" = netzartig. Die [Samen](#) sind oval, an einem Ende rund, am anderen zugespitzt. Im Innern sind sie grün. Ein großer Teil der Samen ist polyembryonisch.

Als erstes genanntes Ursprungsland

☐ Indien ☐ China ☐ England ☐ Japan ☐ Deutschland

Als zweites genanntes Ursprungsland

☐ Indien ☐ China ☐ England ☐ Japan ☐ Deutschland

Das Jahrhundert des ersten Anbaus (gesicherte Annahme)

☐ 15 v. Chr. ☐ 14 v. Chr. ☐ 13 v. Chr. ☐ 12 v. Chr. ☐ 11 v. Chr.

Das Jahrhundert der ersten Einführung in Europa

☐ 17AD ☐ 18AD ☐ 19AD ☐ 20AD ☐ 21AD

Das Land des ersten Anbaus (gesicherte Annahme)

☐ Indien ☐ China ☐ England ☐ Japan ☐ Deutschland

ABBILDUNG 6.3: Teil einer Aufgabe zur Informationsrecherche aus einer Textauswahl von Wikipedia. In diesem Fall wurde ein Zeitlimit als Stressor eingesetzt.

Lehne, eine auf der Sitzfläche des Stuhls befestigt. Diese besaßen eine Matrix aus 8×8 drucksensitiven Zellen, die eine Auflösung von 4,5 cm ermöglichten. Die dritte Matte lag auf dem Boden. Auf diese sollten Studienteilnehmer beim Sitzen ihre Füße platzieren. Sie bestand aus 16×14 Zellen. Die Abtastrate der beiden eingesetzten per USB-Schnittstelle verbundenen Mikrocontroller des Herstellers lag bei 20 Hz. Die Quantisierung der einzelnen Druckzellen besaß eine Auflösung von 12 bit (Werte zwischen 0 und 4095).

Die physiologischen Daten wurden mit einem IOM-Sensor über die USB-Schnittstelle aufgezeichnet, der BVP- sowie EDA-Daten mit 27 Hz lieferte. Die Herzrate wurde aus dem BVP-Signal durch SSI berechnet. Zudem war ein stationärer Eyetracker (The Eye Tribe) mit einer Abtastrate von 30 Hz im Einsatz. Er war per USB-Schnittstelle angeschlossen. Die kapazitiven Berührungsdaten der genutzten Microsoft Touchmouse wurden für den Datensatz ebenfalls aufgezeichnet (Auflösung 15×13 bei 50 Hz), sie wurden allerdings in dieser Arbeit nicht weiter analysiert, da diese stark vom GUI-Design abhängen. Da es sich bei der Touchmouse um eine reguläre Funk-Maus handelte, wurde diese zudem als Eingabegerät genutzt.

Das SSI-Framework wurde genutzt, um die verschiedenen Sensordaten zu synchronisieren, Merkmale zu berechnen und die Daten aufzuzeichnen. Hierfür wurden vom Autor folgende SSI-Plugins (mit-)entwickelt: *sensingtex*, *touchmouse*, *pressuremat* (siehe Anhang C). Alle Eingaben, die beim Stroop-Test oder der Informationsrecherche vom Nutzer ausgeführt wurden, sind mit Zeitstempeln, als SSI-Events, gespeichert worden. In Abbildung 6.5 sind die Anzeigen von SSI dargestellt.

Studienablauf An der Studie nahmen 10 (Frauen: 2, Männer: 8) gesunde Personen mit einem Durchschnittsalter von 31 Jahren teil. Alle Teilnehmer waren Rechtshänder. Daher wurden die Fingerkontakte des IOM-Sensors an der linken Hand befestigt, sodass die rechte Hand für die

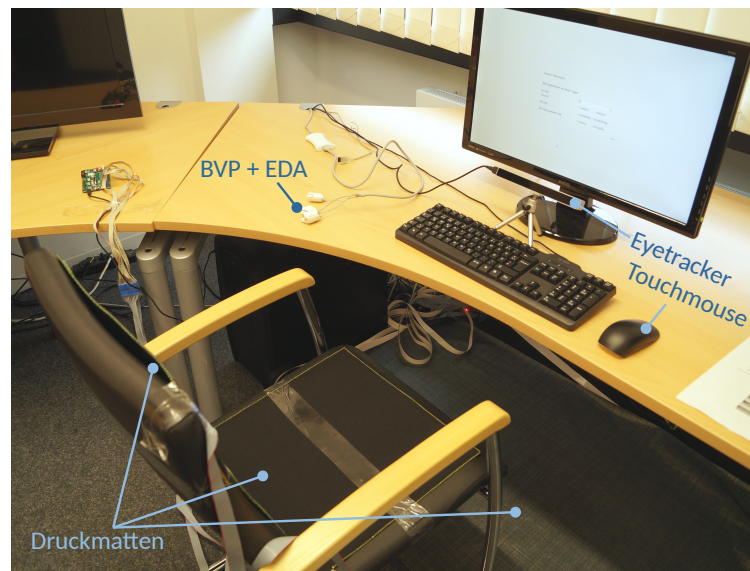
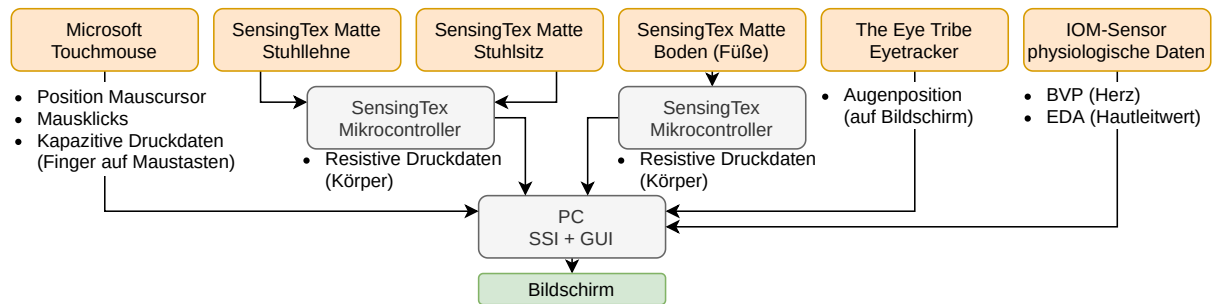


ABBILDUNG 6.4: Studienaufbau: Hardware.

Verwendung der Maus frei war. Die Maus war ausreichend, um das am Bildschirm dargestellte GUI zu bedienen.

Die Teilnehmer wurden in zwei Gruppen mit unterschiedlicher Testreihenfolge aufgeteilt. Die Reihenfolge der Tests war zwischen den Gruppen umgekehrt. Dies sollte mögliche Störeffekte verhindern, die ansonsten durch die Wahl einer bestimmten Reihenfolge der Tests auftreten können. Die Abläufe sind in Abbildung 6.6 dargestellt. Vor jedem Experiment wurden die Teilnehmer über den Ablauf informiert. Daraufhin wurde der IOM-Sensor an drei Fingern einer Hand angeschlossen und der Eyetracker kalibriert. Anschließend wurden folgende demografische Fragen durch die Software abgefragt: Alter, Geschlecht, dominante Hand und die Hand, die bei der Nutzung einer Maus eingesetzt wird. Es folgte ein kurzer Test des kongruenten und inkongruenten Stroop-Test, um zu überprüfen, ob die Erklärungen richtig verstanden wurden. Anschließend begann das Experiment mit einem fünf Minuten langen Entspannungsvideo (Landschaften mit ruhiger Musik). Es folgte ein NASA-TLX-Fragebogen. Dieser bestand aus sechs Fragen, bei denen die mentale, physische und zeitliche Beanspruchung, die Leistung, der Aufwand und die Frustration abgefragt werden. Die Frage nach der physischen Beanspruchung wurde in diesem Fall entfernt, da diese bei Büroarbeitern in der Regel nicht relevant ist. Nun folgte der Studienabschnitt mit den beiden Stroop-Tests. Bei der ersten Gruppe wurde zuerst der inkongruente Test durchgeführt, danach folgte der kongruente. Bei der zweiten Gruppe war dies umgekehrt. Anschließend wurde ein zwei Minuten langes Entspannungsvideo gezeigt und daraufhin sollte ein TLX-Fragebogen ausgefüllt werden. Im Anschluss wurden zwei Aufgaben zur Informationsrecherche gestellt. Eine davon musste unter Zeitdruck durchgeführt werden. Nach jeder Aufgabe folgten ein Entspannungsvideo und TLX-Fragebogen. Die Texte aus denen Informationen herausgefunden werden mussten, bestanden aus 900 Wörtern und stammten aus der deutschen Wikipedia. Die Inhalte bezogen sich auf das neutrale

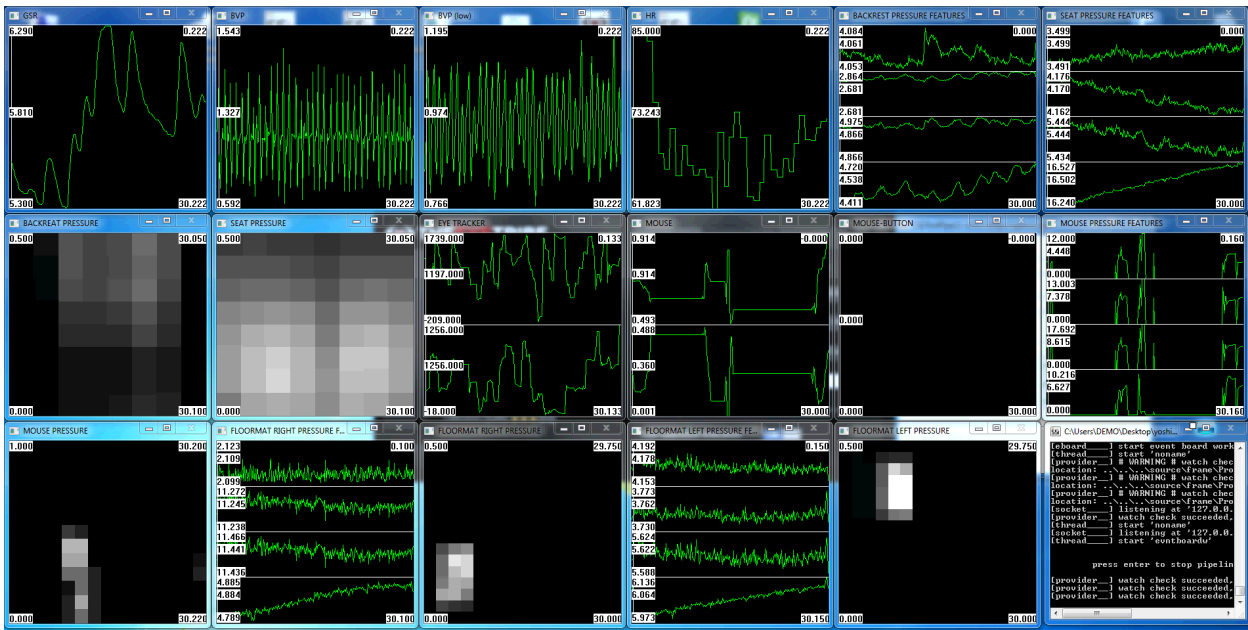


ABBILDUNG 6.5: Anzeigen der Sensordaten und Features von SSI während der Aufnahme.

Thema „Früchte“. Bei der Aufgabe unter Zeitdruck wurde von einer normalen Lesegeschwindigkeit von 200 bis 300 Wörtern pro Minute ausgegangen, die beispielsweise auch in der Studie von Jackson und McClelland (1975) herangezogen wurde. Es wurde daher eine Bearbeitungszeit von drei Minuten festgelegt.

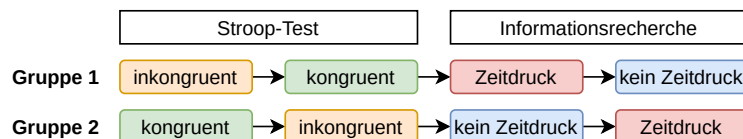


ABBILDUNG 6.6: Studienablauf abhängig von der Gruppe zur Reduktion von Störeffekten, die durch eine bestimmte Reihenfolge ausgelöst werden könnten.

Auswertung

Merkmale Es wurden verschiedene statistische Merkmale aus den Signalen berechnet, die in 10 Sekunden lange Datenfenster aufgeteilt wurden. Jeder Stimulus dauerte 180 Sekunden.

Bei den **Druckmatten** wurde der „Center of Pressure“ berechnet (Beispiel siehe Abschnitt 3.4), der den gewichteten Mittelpunkt darstellt. Die x- und y-Achse des berechneten Punkts entsprachen jeweils einem Merkmal. Zusätzlich wurde die euklidische Distanz berechnet, um ein kombiniertes Merkmal zu erhalten. Die Gesamtintensität, die eine Summe der Werte aller Druckzellen einer Matte darstellt, wurde ebenfalls als Merkmal berechnet. Für die beschriebenen vier Merkmale wurden sechs typische statistische Funktionen, Mittelwert, Standardabweichung, erste und zweite Differenz und die normalisierte erste und zweite Differenz, berechnet. Zusätzlich wurden frequenzbasierte Merkmale einbezogen, die in den Wertebereichen 0,1-0,7 Hz, 0,7-1,3 Hz (Peterka & Loughlin, 2004) und 3,5-8,0 Hz (Krafczyk et al., 1999) lagen. Laut den angegebenen Quellen besitzen Körperschwankungen in diesen Frequenzbereichen einen Bezug zu Stress und Emotionen. Für die Auswertungen wurde zusätzlich das Frequenzband dazwischen (1,3-3,5 Hz) berücksichtigt.

Bei den **physiologischen Sensoren** wurden ebenfalls die sechs statistischen Funktionen auf die Signale angewandt und zusätzlich beim Hautleitwert, wie in der Arbeit von Zhai et al. (2005), die „peak response time“, „peak amplitude“ und Energie berechnet. Aus dem BVP-Signal wurden

die mittlere Amplitude, Schiefe (engl. skewness) und Kurtosis, die auch in der Publikation von Mokhayeri et al. (2011) eingesetzt wurden, zusätzlich zu den sechs anderen statistischen Merkmalen berechnet. Für die aus dem BVP-Signal berechnete Herzrate wurden wiederum die sechs statistischen Merkmale berechnet. Wie in der Arbeit von Koldijk et al., 2014 wurde der Parameter der Herzratenvariabilität „RMSSD“ (root mean square successive difference), in diesem Fall aus dem BVP-Signal, bestimmt. Als weitere Merkmale wurden die Frequenzbänder in den Bereichen $f_a = 0,05-0,15$ Hz und $f_b = 0,15-0,4$ Hz sowie deren Verhältnis (f_b/f_a) zueinander mit dem BVP-Signal ermittelt.

Aus den Daten der **Augenposition**, wurden das Blinzeln, Fixierungen, Sakkaden (Bewegungen zwischen zwei Fixierungen) und Suchvorgänge als Merkmale verwendet, die auch in den Arbeiten von Manor und Gordon (2003) und Nodine et al. (1992) genutzt wurden. Hieraus wurden jeweils die Zeitdauer, die räumliche Distanz (ohne Blinzeln) und die Anzahl des Auftretens bestimmt. Aus diesen Werten wurden das Maximum, Minimum und die Summe ermittelt.

Klassifikation Vor dem Training der Klassifikationsmodelle wurde eine Selektion der Merkmale (engl. Feature Selection) für jedes Signal (EDA, BVP, HR, Druckverteilungen (Boden, Stuhllehne, Sitz) und Eyetracker) mittels „sequential backward selection“ (SBS) durchgeführt. Es wurden hierbei auch Kombinationen von Signalen getestet: physiologische Signale (EDA + BVP + HR), Druckmatten (Boden + Lehne + Sitz) und Eyetracker-Daten (= Augen).

Es wurden nach der Merkmalsselektion die „Modality Level Classification“ sowie „Decision Level Classification“ getestet, die in (J. Kim, 2007) für Emotionserkennung eingesetzt wurden.

Bei der **Modality Level Classification** wurden pro Modalität und Zustand (entspannt, konzentriert, gestresst) Trefferquoten mittels pLDA (pseudo Linear Discriminant Analysis) berechnet. Das Ergebnis jeder Klassifikation für jede Modalität und jeden Zustand wurde als „Entscheidung“ bezeichnet. Zusätzlich wurde als Referenz die Trefferquote über alle Modalitäten bestimmt.

Bei der **Decision Level Classification** wurde eine Klassifikation durchgeführt, indem mehrere Entscheidungen mehrerer Modalitäten kombiniert wurden. Hierfür wurden gewichtete Entscheidungen (X_j) für jeden Zustand j berechnet (siehe Formel 6.1). Die Erkennungsrate für Zustand j basierend auf den Merkmalen der Modalität i wurde als P_{ij} definiert. Die Entscheidung für Zustand j basierend auf der Modalität i wurde als X_{ij} definiert.

$$X_j = \frac{\sum_i P_{ij} X_{ij}}{\sum_i P_{ij}} \quad (6.1)$$

Die gewichteten Entscheidungen können als Merkmale verwendet werden. Diese entscheidungsbasierten Merkmale (engl. decision level features) wurden eingesetzt, um fünf Klassifikatoren zu trainieren: kNN (wobei $k = 1$), pLDA, SVM mit linearem Kernel, SVM mit RBF- (radial basis function) Kernel und Fuzzy Logic.

Ergebnisse

Die Trefferquoten der **Modality Level Classification** mit pLDA-Klassifikator sind in Tabelle 6.1 dargestellt. Es sind hierbei die beiden Stress-Stimuli, verschiedenen Modalitäten sowie einige Kombinationen von Modalitäten („Physiolog.“ = physiologisch, „Druckmat.“ = Druckmatten und „Alle“) angegeben. Bei „personenunabhängigen“ Erkennungsraten wurden für die Trainings- und Test-Daten, Daten verschiedener Studienteilnehmer verwendet, bei „personenabhängigen“ Raten stammten die Daten von derselben Person. Für die Bestimmung der Trefferquoten wurde die Leave-One-Out-Kreuzvalidierung eingesetzt. Zusätzlich werden die Konfusionsmatrizen der Eyetracker-Signale (Augen), physiologischen Signale (EDA + BVP + HR) sowie der Druckmatten (Boden + Lehne + Sitz) in Tabelle 6.2 angegeben. Die mittleren Trefferquoten entsprechen der Treffergenauigkeit, die in einigen anderen Studien ermittelt wurden.

Stroop-Test								
Modalität	personenunabhängig				personenabhängig			
	entspannt	konzentriert	gestresst	Mittelwert	entspannt	konzentriert	gestresst	Mittelwert
EDA	76,7%	22,8%	28,9%	42,8%	100,0%	97,2%	95,6%	97,6%
BVP	42,2%	52,8%	36,7%	43,9%	77,2%	98,3%	62,8%	79,4%
HR	36,7%	26,7%	50,6%	38,0%	88,9%	83,9%	57,8%	76,9%
Boden	25,0%	35,6%	46,1%	35,6%	82,8%	74,4%	78,9%	78,7%
Lehne	62,8%	31,1%	29,4%	41,1%	100,0%	87,2%	82,2%	89,8%
Sitz	38,9%	32,8%	53,9%	41,9%	100,0%	99,4%	93,9%	97,8%
Augen	66,1%	53,3%	21,1%	46,9%	95,0%	65,6%	63,9%	74,8%
Physiolog.	57,2%	40,0%	29,4%	42,2%	98,3%	98,9%	93,9%	96,9%
Druckmat.	45,6%	58,3%	46,1%	50,0%	99,4%	89,4%	85,6%	90,2%
Alle	70,0%	18,9%	60,6%	53,3%	100,0%	92,2%	93,9%	94,3%

Informationsrecherche								
Modalität	personenunabhängig				personenabhängig			
	entspannt	konzentriert	gestresst	Mittelwert	entspannt	konzentriert	gestresst	Mittelwert
EDA	64,4%	10,6%	47,8%	40,9%	100,0%	97,7%	100,0%	99,1%
BVP	47,8%	20,0%	49,4%	39,1%	90,3%	95,4%	94,0%	93,0%
HR	12,8%	38,3%	50,0%	33,7%	92,6%	63,9%	72,7%	75,7%
Boden	46,1%	52,8%	30,6%	43,1%	83,8%	84,7%	87,0%	84,1%
Lehne	62,8%	30,6%	31,1%	41,5%	94,4%	78,7%	73,6%	82,8%
Sitz	42,2%	27,2%	52,8%	40,7%	96,8%	87,5%	91,2%	92,4%
Augen	46,1%	48,3%	40,6%	45,0%	87,0%	59,3%	61,6%	70,6%
Physiolog.	58,3%	32,8%	34,4%	41,9%	98,6%	95,8%	98,2%	97,0%
Druckmat.	73,9%	55,0%	49,4%	57,8%	86,6%	81,9%	81,5%	91,9%
Alle	80,6%	55,0%	32,2%	53,7%	96,3%	88,0%	87,5%	91,7%

TABELLE 6.1: Übersicht über die Trefferquoten.

Bei beiden Stress-Stimuli (Stroop-Test bzw. Informationsrecherche) zeigten in den personenunabhängigen Fällen die Klassifikationen über die kombinierten Druckmatten-Signale höhere Trefferquoten (50,0 % (Stroop) bzw. 57,8 %) als über die Eyetracker-Signale (46,9 % (Stroop) bzw. 45,0 %) und physiologischen Signale (42,2 % (Stroop) bzw. 41,9 %). Bei den personenabhängigen Ergebnissen schnitten allerdings die physiologischen Signale (96,9 % (Stroop) bzw. 97,0 %) besser ab als die Druckmatten (90,2 % (Stroop) bzw. 91,9 %) und Eyetracker-Daten (74,8 % (Stroop) bzw. 70,6 %). Wie bei den Trefferquoten sind die Genauigkeiten in den personenunabhängigen Fällen bei Druckmattendaten (51,6 % (Stroop) bzw. 60,5 %) höher als bei den Eyetracker- (45,2 % (Stroop) bzw. 45,2 %) und Physiologischen- (41,9 % (Stroop) bzw. 41,9 %) Signalen (siehe Tabelle 6.2).

Die Konfusionsmatrizen zeigen bei den Druckmatten- und Eyetracker-Daten im personenunabhängigen Fall und beiden Stimuli, dass der Zustand „gestresst“ häufig mit „konzentriert“ verwechselt wurde. Bei den physiologischen Daten war dies allerdings nicht so deutlich. Die Klassen „konzentriert“ und „gestresst“ scheinen mit den aufgezeichneten Daten schwieriger unterscheidbar gewesen zu sein. Es existierte höchstwahrscheinlich ein personenabhängiger, fließender Übergang zwischen diesen Zuständen, da bei der personenabhängigen Klassifizierung diese Verwechslungen wesentlich seltener auftraten.

Bei der **Decision Level Classification** wurden mit der Decision-Level-Fusion (verwendete entscheidungsbasierte Merkmale gemäß Formel 6.1) die drei Stress-Zustände mit verschiedenen Klassifikatoren für die beiden Stimuli sowie personenabhängig und personenunabhängig klassifiziert. Die Ergebnisse sind in Tabelle 6.3 dargestellt. Die Ergebnisse des pLDA-Klassifikators nach der „Feature-Level-Fusion“, bei der alle Merkmale aller Modalitäten kombiniert (fusioniert) wurden, sind ebenfalls als Referenz angegeben (entsprechen den Zeilen mit der fusionierten Modalität „Alle“ und den Spalten „Mittelwert“ in Tabelle 6.1). Für die Informationsrecherche zeigte die Decision-Level-Fusion höhere Trefferquoten als die Feature-Level-Fusion. Beim Stroop-Test ist dies allerdings nicht der Fall.

Stroop-Test										
Augen	personenunabhängig					personenabhängig				
	vorhergesagter Zustand			Trefferquote	vorhergesagter Zustand			Trefferquote		
	entspannt	konzentriert	gestresst		entspannt	konzentriert	gestresst			
entspannt	119	39	22	66,1%	46,9%	171	1	8	95,0%	74,8%
konzentriert	45	96	39	53,3%		4	118	58	65,6%	
gestresst	65	77	38	21,1%		7	58	115	63,9%	
Genauigkeit	52,0%	45,3%	38,4%			94,0%	66,7%	63,5%		
	45,2%					74,7%				

Informationsrecherche										
Augen	personenunabhängig					personenabhängig				
	vorhergesagter Zustand			Trefferquote	vorhergesagter Zustand			Trefferquote		
	entspannt	konzentriert	gestresst		entspannt	konzentriert	gestresst			
entspannt	83	52	45	46,1%	45,0%	168	9	3	93,3%	70,6%
konzentriert	37	87	56	48,3%		14	94	72	52,2%	
gestresst	45	62	73	40,6%		5	56	119	66,1%	
Genauigkeit	50,3%	43,3%	42,0%			89,8%	59,1%	61,3%		
	45,2%					70,1%				

Stroop-Test										
Physiologisch	personenunabhängig					personenabhängig				
	vorhergesagter Zustand			Trefferquote	vorhergesagter Zustand			Trefferquote		
	entspannt	konzentriert	gestresst		entspannt	konzentriert	gestresst			
entspannt	103	42	35	57,2%	42,2%	177	0	3	98,3%	97,0%
konzentriert	67	72	41	40,0%		1	178	1	98,9%	
gestresst	64	63	53	29,4%		3	8	169	93,9%	
Genauigkeit	44,0%	40,7%	41,1%			97,8%	95,7%	97,7%		
	41,9%					97,1%				

Informationsrecherche										
Physiologisch	personenunabhängig					personenabhängig				
	vorhergesagter Zustand			Trefferquote	vorhergesagter Zustand			Trefferquote		
	entspannt	konzentriert	gestresst		entspannt	konzentriert	gestresst			
entspannt	105	42	33	58,3%	41,9%	177	3	0	98,3%	97,0%
konzentriert	76	59	45	32,8%		3	171	6	95,0%	
gestresst	69	49	62	34,4%		0	4	176	97,8%	
Genauigkeit	42,0%	39,3%	44,3%			98,3%	96,1%	96,7%		
	41,9%					97,0%				

Stroop-Test										
Druckmatten	personenunabhängig					personenabhängig				
	vorhergesagter Zustand			Trefferquote	vorhergesagter Zustand			Trefferquote		
	entspannt	konzentriert	gestresst		entspannt	konzentriert	gestresst			
entspannt	82	43	55	45,6%	50,0%	179	0	1	99,4%	91,5%
konzentriert	21	105	54	58,3%		2	161	17	89,4%	
gestresst	26	71	83	46,1%		6	20	154	85,6%	
Genauigkeit	63,6%	47,9%	43,2%			95,7%	89,0%	89,5%		
	51,6%					91,4%				

Informationsrecherche										
Druckmatten	personenunabhängig					personenabhängig				
	vorhergesagter Zustand			Trefferquote	vorhergesagter Zustand			Trefferquote		
	entspannt	konzentriert	gestresst		entspannt	konzentriert	gestresst			
entspannt	133	38	9	73,9%	59,4%	165	10	5	91,7%	84,1%
konzentriert	43	99	38	55,0%		6	145	29	80,6%	
gestresst	16	75	89	49,4%		2	34	144	80,0%	
Genauigkeit	69,3%	46,7%	65,4%			95,4%	76,7%	80,9%		
	60,5%					84,3%				

TABELLE 6.2: Konfusionsmatrizen für die Klassifikationen des Stresslevels mittels Eyetracker- (Augen), physiologischen (EDA + BVP + HR) und Druckmattendaten (Boden + Lehne + Sitz).

		Stroop-Test		Informationsrecherche	
		personen-unabhängig	personen-abhängig	personen-unabhängig	personen-abhängig
Decision Level Fusion	kNN	45,74%	81,67%	49,81%	90,93%
	pLDA	48,89%	85,74%	51,48%	92,96%
	lineare SVM	46,48%	84,07%	49,81%	93,15%
	RBF SVM	45,56%	84,26%	49,26%	93,52%
	Fuzzy Logic	45,93%	80,00%	54,26%	87,59%
Feature Level Fusion		53,33%	94,26%	53,70%	91,67%

TABELLE 6.3: Trefferquoten für Decision-Level-Fusion mit verschiedenen Klassifikatoren bzw. der Feature-Level-Fusion.

Insgesamt ist eines der interessantesten Ergebnisse der Auswertungen, dass im personenunabhängigen Fall für beide Stressoren (Stroop-Test und Informationsrecherche) die Klassifikation mit den kombinierten Daten aller Druckmatten im Mittel eine höhere Trefferquote über alle Zustände zeigte, als die Klassifikationen mit der Kombination der physiologischen Daten oder mit den Signalen des Eyetrackers. Dies widersprach den Erwartungen, da die physiologischen Daten und Augenbewegungen als unwillkürliche Modalitäten betrachtet werden und direkt neurale Aktivitäten, zu denen auch Emotionen und Stress zählen, reflektieren. Dem gegenüber kann die Körperhaltung willkürlich beeinflusst werden. Weitere Untersuchungen wären hierbei notwendig, um dieses Ergebnis zu überprüfen.

Die meisten in der Literatur auffindbaren Studien zur Stresserkennung beinhalten keine Druckmatten, sondern primär physiologische Sensoren für Hautleitwert (EDA) oder Daten des Herzens (Herzrate (HR), Herzratenvariabilität (HRV)), die mittels Blutvolumenpuls oder EKG gemessen werden. Auch der Literaturüberblick von Akbar et al. (2019) bestätigt dies. Mit empfindlicheren Druckmatten wäre es allerdings auch bedingt möglich, die physiologischen Daten Herzrate und Atmung aus geringsten Druckänderungen (entspricht einem Ballistokardiogramm) zu gewinnen (Sadek & Biswas, 2019) und somit diese extrahierten Biosignale, anstelle der Druckmattendaten und Merkmale, zu nutzen.

Ein Vergleich der Erkennungsraten des Stresslevels mit anderen Studien ist nur begrenzt möglich, da diese teilweise zur Vereinfachung nur zwischen zwei Zuständen unterschieden. Kaum vergleichbar ist auch der Ansatz, bei dem ein Unterschied zum Grundzustand (entspannt) gemessen wird, wobei ab einer bestimmten Abweichung Stress angenommen wird. Zusätzlich wurden häufig deutlich verschiedene Studiendesigns und Stress-Stimuli eingesetzt. Falls mehrere Modalitäten verwendet wurden, sind diese zudem nicht immer getrennt analysiert worden.

Ein eingeschränkter Vergleich der eigenen Arbeit ist mit den Erkennungsraten aus der Studie von Zhai und Barreto (2006) möglich, bei der jedoch ausschließlich zwischen konzentriertem Zustand (kongruenter Stroop-Test) und Stresszustand (inkongruenter Stroop-Test) unterschieden wurde. Es handelt sich daher nur um ein Zwei- (zufällige Erkennungsrate: 50 %) und nicht um ein Drei-Klassen-Problem (zufällige Erkennungsrate: ca. 33,3 %), das in der eigenen Arbeit untersucht wurde. An der Studie von Zhai und Barreto (2006) nahmen 32 Studienteilnehmer teil. Von den Teilnehmern wurden die physiologischen Signale BVP, EDA, ST (Skin Temperature = Hauttemperatur) und PD (Pupil Diameter = Pupillendurchmesser) aufgezeichnet. Mit einem SVM-Klassifikator und 20-facher Kreuzvalidierung konnte eine Treffergenauigkeit von 90,10 % erreicht werden. Die Stress-Stimuli sind in diesem Fall identisch zu der vorliegenden Arbeit, die physiologischen Signale unterscheiden sich allerdings. In der Arbeit von Zhai und Barreto (2006) wurden einzelne Modalitäten testweise aus der Klassifikation entfernt und die daraus resultierenden Treffergenauigkeiten gemessen. Wenn der Pupillendurchmesser wie bei der eigenen Arbeit nicht verwendet wurde, sank die Treffergenauigkeit auf 61,5 %. Die Treffergenauigkeit der eigenen Untersuchung, basierend auf der Feature-Level-Fusion mit physiologischen Daten und vergleichbaren Stimuli für die Zustände „entspannt“, „konzentriert“ und „gestresst“, lag bei 41,9 % (siehe Tabelle 6.2 Physiologisch, Stroop-Test, personenunabhängig; mittlere Trefferquoten = Treffergenauigkeit). Der deutliche Unterschied

bei der Treffergenauigkeit kann folgende Ursachen besitzen: Zhai und Barreto (2006) lösten nur ein für Klassifikatoren einfacheres Zwei-Klassen-Problem, sie setzten zusätzliche Biosignale ein, sowohl die Datenmenge als auch die Methode der Validierung unterscheiden sich von der eigenen Arbeit.

Eine der wenigen Stress-Studien mit Nutzung einer Druckmatte findet sich in der Arbeit von Arnrich et al. (2009). In dieser konnte mit Daten von 28 Studienteilnehmern und einem Kohonen-Netz als Klassifikator eine Treffergenauigkeit von 73,8 % zwischen einem konzentrierten Zustand und Stresszustand mit einer Druckmatte (Sitz) mit 1024 Sensorelementen, personenunabhängig erreicht werden. Es wurden verschiedene Stress-Stimuli für den Stresszustand eingesetzt, die sozialen und kognitiven Stress verursachten. Es wurden im Unterschied zur eigenen Arbeit nur zwei Zustände betrachtet und ausgewertet. Ein direkter Vergleich der Erkennungsraten mit den eigenen Ergebnissen ist damit nicht sinnvoll, zudem die Druckmatten bei der Untersuchung von Arnrich et al. (2009) deutlich höhere Auflösungen besaßen.

Zusammenfassung

In dieser Arbeit konnte gezeigt werden, dass mit verschiedenen Modalitäten Stress erkannt werden kann. Hierbei wurde insbesondere die Nutzung der Daten von drucksensitiven Matten im Vergleich zu den klassischen physiologischen Signalen sowie den Daten eines Eyetrackers analysiert, die bislang noch nicht in diesem Umfang untersucht worden waren. Die Ergebnisse der Analyse legen nahe, dass selbst Matten mit relativ niedriger Auflösung für eine Stresserkennung hinreichend gute Daten liefern können. Die personenunabhängigen Trefferquoten lagen bei den beiden Stimuli bei 53,3 % bzw. 54,3 %, was deutlich über der zufälligen Erkennungsrate von 33,3 % bei drei Klassen liegt. Im personenabhängigen Fall lagen sie bei ca. 94,3 % bzw. 93,5 %.

Druckmatten könnten für eine Stresserkennung in Stühle integriert bzw. auf dem Boden platziert werden und daher in einigen Anwendungsbereichen eine Alternative zu tragbarer Sensorik darstellen. Für weitere Untersuchungen wäre allerdings ein umfangreicherer Datensatz notwendig. Dieser könnte auch zu einer besseren personenunabhängigen Erkennung beitragen. Problematisch ist, dass drucksensitive Matten derzeit nur wenig verbreitet sind und sich teilweise stark durch die eingesetzten Messprinzipien und technischen Umsetzungen (Auflösung / Abtastraten / Empfindlichkeit) unterscheiden. Der in dieser Arbeit vorgestellte Studienaufbau wäre allerdings generell für eine kontrollierte Datenaufzeichnung geeignet. Ergänzend könnten Feldstudien mit selbstbewertetem Stresslevel Daten aus dem Alltag von Büroarbeitern liefern.

Zusätzlich zur Erkennung von Stress, sind Druckmatten im Büroalltag auch geeignet, um beispielsweise regelmäßig auf sogenannte Mikropausen mit Atmungsübungen zur Erholung hinzuweisen und diese zu überwachen (Ren et al., 2019). Hierdurch kann der Stresslevel reduziert werden. Untersuchungen bezüglich der Kombination einer Stresserkennung sowie dem Abbau von Stress mittels geeigneter Übungen wären daher weiterführend möglich.

6.3.2 Erfassung von Lachen mit Smartphones

Einleitung

Lachen stellt als positive Emotion eine bedeutsame Information für das Wohlbefinden dar. Im Bezug auf eine Gruppe kann es zudem als Bestandteil einer sozialen Interaktion dienen. In dieser Arbeit wurde untersucht, wie mittels multimodaler Signalverarbeitung eine Echtzeiterkennung von Lachen auf mobilen Geräten durchgeführt werden kann. Im Rahmen der Arbeit wurde zunächst ein Datenkorpus bestehend aus Audio- und Bewegungsdaten mit mobileSSI (siehe Abschnitt 4.1.2) und Smartphones aufgezeichnet, das zum Trainieren eines Modells verwendet wurde. Mit mobileSSI wurde eine Echtzeitklassifizierungs-Pipeline erstellt, die auf einem Smartphone ausgeführt werden konnte. Die Ergebnisse dieses Projektes wurden in den eigenen Publikationen (Flutura, Wagner,

Lingenfelser, Seiderer & André, 2016b), (Flutura, Wagner, Lingenfelser, Seiderer & André, 2016c) sowie Flutura, Wagner, Lingenfelser, Seiderer und André (2016a) veröffentlicht.

Datenaufzeichnung

Für das Training der Klassifizierungsmodelle wurden beschriftete Daten benötigt. Diese wurden nicht im Labor, sondern in einem alltäglichen Szenario mit mobileSSI aufgezeichnet und anschließend manuell annotiert.

Die Daten wurden von drei Personen in einem Gasthaus aufgezeichnet, da sich hier Leute oft treffen und fröhliche Gespräche führen. Die Studienteilnehmer waren miteinander bereits bekannt und konnten das Gespräch frei führen. Sie trugen hierbei jeweils ein Samsung Galaxy S4 Smartphone (Android 5.0.1) in einer Brusttasche an dem ein Clip-Mikrofon angeschlossen war, das am Kleidungsstück am Halsausschnitt befestigt wurde. Die Audiodaten wurden mit einer Abtast-rate von 16 kHz, die Beschleunigungsdaten mit 100 Hz aufgezeichnet. In zwei Sitzungen wurden insgesamt ca. 3,5 Stunden Daten pro Person erfasst. Zur Erstellung eines Echtzeitklassifikators musste eine Datensegmentierung durchgeführt werden. Ein Datenfenster war 1 Sekunde lang und es wurde eine Überlappung von 0,4 Sekunden genutzt. Hierdurch entstanden 21500 Samples, von denen 875 Lachen enthielten.

Im Unterschied zu anderen Korpora, wie beispielsweise von Jansen et al. (2018), bei dem spontanes Lachen in Laborumgebung u. a. mit Audio- und Accelerometer-Daten aufgezeichnet wurden, waren bei dem in dieser Arbeit erstellten Korpus Sensordaten von Smartphones enthalten, die in einer Alltagsumgebung aufgezeichnet wurden.

Merkmalsextraktion, Klassifizierung und Evaluation

Für die Audiodaten wurden Merkmale berechnet, die von dem in SSI integrierten EmoVoice-Framework (Vogt et al., 2008) stammten. Hieraus entstanden Merkmalsvektoren mit 1451 Werten. In diesen waren beispielsweise auch die besonders in der Spracherkennung häufig eingesetzten MFCCs (Mel Frequency Cepstral Coefficients³) enthalten. Für die Beschleunigungsdaten wurden neun Merkmale (mean, standard deviation, min, max, range, zero crossing rate, peak count, pulse rate, energy) über die Rohdaten sowie deren erste und zweite Ableitung von jeder Achse berechnet. Insgesamt resultierte dies in Merkmalsvektoren aus 81 Werten.

Die Evaluation wurde über die Daten der zwei Sitzungen durchgeführt. In beiden Sitzungen waren die Daten derselben drei Personen aufgezeichnet worden. Es wurden daher die Daten von zwei Personen für das Training der Modelle verwendet und die Daten der dritten Person für das Testen. Diese Aufteilung war fix und entsprach der, die in der Arbeit von Lingenfelser et al. (2014) eingesetzt wurde, da die Ergebnisse vergleichbar sein sollten.

Zunächst wurden die beiden Modalitäten getrennt genutzt, um SVM-Modelle mit zwei Klassen zu erstellen. Die Klassen waren „Lachen“ und „Nicht-Lachen“. Zum Ausgleich der Datenmenge pro Klasse wurde „random under-sampling“ eingesetzt. Mit den Merkmalen der Audiodaten konnten im ungewichteten Mittel (engl. unweighted average = UA) eine Treffergenauigkeit von 81,5 %, mit den Bewegungsdaten 72,2 % erreicht werden (siehe Tabelle 6.4). Die Ergebnisse hinsichtlich der Audiodaten können mit denen aus der Arbeit von Lingenfelser et al. (2014) verglichen werden. Hierbei wurden Audio- und Videodaten unter Laborbedingungen aufgezeichnet. Die Auswertung fand dabei identisch statt. Es wurde bei den Audiodaten, unter Laborbedingungen, eine ungewichtete mittlere Treffergenauigkeit von 90 % erreicht. Somit lag diese 8,5 % höher als bei den Daten aus dem Gasthaus. Die Reduktion der Treffergenauigkeit bei der eigenen Untersuchung dürfte vor allem durch Umgebungsgeräusche im Wirtshaus verursacht worden sein. Die Videodaten waren

³Berechnung: Spektrum (FFT) → Logarithmus vom Amplitudenspektrum → Anwendung Mel-Filterbank (Reduktion der Frequenzbänder: höhere Frequenzen werden vergleichbar der menschlichen Wahrnehmung stärker zusammengefasst) → Anwendung diskrete Kosinustransformation (Logan, 2000)

nicht direkt mit den Bewegungsdaten der eigenen Studie vergleichbar, da mithilfe der Videos nicht nur die Körperbewegungen, sondern auch Gesichtsausdrücke, berücksichtigt wurden.

	Unimodale Klassifikation			Multimodale Klassifikation	
	Accelerometer	Audio		Decision Fusion	Event Fusion
Lachen	80,95 %	76,19 %	Lachen	78,57 %	83,33 %
¬ Lachen	63,42 %	86,70 %	¬ Lachen	86,62 %	85,95 %
Mittelwert	72,19 %	81,45 %	Mittelwert	82,59 %	84,64 %

TABELLE 6.4: Ergebnisse pro Modalität (links) und nach der Decision bzw. Event Fusion (rechts).

Zusätzlich wurden zwei Methoden zur Fusion der Klassifikationsergebnisse untersucht. Diese waren die „Decision Fusion“ und „Event Fusion“. Bei der entscheidungsbasierten Fusion (Decision Fusion) werden die Ergebnisse mehrerer Klassifikatoren genutzt, um die Klassen final zu bestimmen. In diesem Fall wurden die Ergebnisse mit der Produktregel berechnet. Die ereignisbasierte Fusion wurde bereits in Lingenfelder et al. (2014) für die multimodale Erkennung von Lachen genutzt und zeigte hierbei gute Ergebnisse. Bei der Erkennung von Lachen ist sie gut geeignet, da mehrere Modalitäten, wie Audio- und Bewegungsdaten genutzt werden können, die allerdings nicht zwingend zeitgleich auftreten und ihre Relevanz gewichtet werden kann. Insbesondere für mobile Anwendungen ist das Verfahren geeignet, da hierbei nicht immer alle Modalitäten zur Verfügung stehen müssen sowie deren aktuelle und generelle Zuverlässigkeit einbezogen werden kann. Beispielsweise stünde die Modalität „Körperbewegung“ bei der Lacherkennung während des Laufens nicht zur Verfügung oder wäre nur sehr unzuverlässig. Umgekehrt ist in geräuschvoller Umgebung das Audiosignal nur noch begrenzt nutzbar.

Bei der ereignisbasierte Fusion werden wie bei der entscheidungsbasierten Fusion mehrere Klassifikatoren eingesetzt. Diese liefern ihre Ergebnisse, sobald diese verfügbar sind, als Event an einen Fusionsalgorithmus. In diesem Fall wurde der in (Lingenfelder et al., 2014) beschriebene Fusionsalgorithmus eingesetzt. Hierbei werden die Ereignisse vergleichbar mit der Masse und Gravitation physischer Objekte behandelt, um hieraus die Wahrscheinlichkeit für die Klasse zu berechnen. Es musste in diesem Fall nur zwischen „Nicht-Lachen“ und „Lachen“ entschieden werden (binäre Klassifikation). Visuell ist die Klasse durch die y-Achse darstellbar (Lach-Wahrscheinlichkeit 0 bis 1) und die Zeit durch die x-Achse. Die Ereignisse besitzen, entsprechend ihrer Zuverlässigkeit eine bestimmte Masse. Die Position auf der y-Achse hängt von der Ausgabe des entsprechenden Klassifikators ab. Über die Zeit werden die Massen verringert (decay). Zur Bestimmung der aktuellen Klasse wird zunächst der Massenschwerpunkt zum Zeitpunkt t auf der y-Achse bestimmt, der von den vergangenen Ereignissen entsprechend ihrer Masse zum Zeitpunkt t und Position abhängt. Anschließend wird ein Schwellwert angewandt.

Bei der „Decision Fusion“ konnte eine ungewichtete mittlere Treffergenauigkeit von 82,6 %, bei der „Event Fusion“ 84,6 % erreicht werden (siehe Tabelle 6.4). Insgesamt verbesserten beide Fusions-Verfahren die unimodalen Ergebnisse. Die „Event Fusion“ schnitt insgesamt am besten ab. Verschiedene Gewichtungen und Abnahmewerte wurden getestet. Eine höhere Gewichtung der Audiodaten (0,7 oder 1,0) führte zu diesem besten Ergebnis. Die Accelerometer-Daten waren hierbei zwischen 0,1 und 0,3 gewichtet. Die höhere Gewichtung der Audiodaten lässt sich auch anhand des besseren Ergebnisses bei den unimodalen Klassifizierungen erklären. Die zeitliche Abnahme des Einflusses (decay) lag bei Audiodaten zwischen 0,6 und 1,0 und bei den Beschleunigungsdaten zwischen 0,5 und 1,5.

Demonstration eines Prototyps mit GUI und Netzwerkkommunikation

Für einen praktischen Test wurde die entwickelte Echtzeiterkennung auf drei Smartphones mit mobileSSI ausgeführt, die wie bei der Datenaufzeichnung drei Personen gegeben wurden. Zusätzlich

wurden die Klassifikationsergebnisse an ein Tablet per Netzwerk gesendet, das ebenfalls mobileSSI ausführte. Das Tablet stellte eine Visualisierung mittels einer Website dar (Kommunikation per WebSocket), bei der ersichtlich war, wer gerade gelacht hat. Die Echtzeiterkennung war mit den Galaxy S4 Smartphones über sieben Stunden hinweg möglich.

Auch zur Untersuchung der sozialen Verbundenheit der Nutzer wäre dieses System geeignet, indem erfasst wird, wie sich Lachen auf andere Personen überträgt. Der Aufbau könnte hierbei, wie bei der Datenaufzeichnung an einem geeigneten Ort, wie beispielsweise einem Gasthaus, in einer Feldstudie eingesetzt werden. Der Einfluss einer Echtzeitvisualisierung auf das Verhalten der Probanden könnte ebenfalls analysiert werden, da es gemeinsames Lachen stimulieren könnte. Anwendungen im Lachtherapie-Bereich wären daher ebenfalls vorstellbar.

Aus technischer Sicht zeigte dieser Prototyp, dass mit mobileSSI und entsprechenden Plugins eine Kommunikation zwischen mehreren mobilen Geräten möglich ist. Für die Darstellung der GUI wurde das vom Autor entwickelte WebSocket-Plugin (siehe Plugin-Übersicht im Anhang C) mit einer HTML5-basierten Website genutzt, die somit eine Systemausgabe erlaubte. Das WebSocket-Plugin ist jedoch zusätzlich in der Lage, Eingaben vom Nutzer an (mobile)SSI zu senden und erlaubt somit die Nutzung moderner GUIs auf PCs und mobilen Geräten. Das Plugin wurde daher auch in einem späteren mobilen Prototyp (siehe Abschnitt 7.3.2) sowie mit SSI beim CARE-Projekt (siehe Abschnitt 9.5 und 9.6), eingesetzt. Zusätzliche Anbindungen an IoT-Hardware und Software sind mit den ebenfalls vom Autor entwickelten Plugins für das MQTT- sowie XMPP-Protokoll möglich und wurden beim CARE-Projekt mit SSI genutzt (siehe Abschnitt 9.5).

Diskussion

Durch den Einsatz der ereignisbasierten Fusion konnte die Treffergenauigkeit verbessert werden und näherte sich durch dieses Verfahren den Ergebnissen von Lingenfelser et al. (2014) bei denen allerdings Daten, unter Laborbedingungen eingesetzt wurden.

In der Arbeit von Fukumoto et al. (2013) konnte unter Laborbedingungen eine Erkennungsrate von 89,2 % bei der Erkennung von Lachen und Lächeln mit einer speziellen Brille mit Foto-Unterbrechern erreicht werden. In diesem Fall wurden bei der Evaluation allerdings ausschließlich Videos zur Stimulation von Lächeln / Lachen eingesetzt und nicht wie in der vorliegenden Arbeit realistische soziale Interaktion über mehrere Stunden durchgeführt. Zudem wurden in (Fukumoto et al., 2013) die Daten auf einem PC in Echtzeit verarbeitet und nicht auf einem mobilen Gerät.

Zusammenfassung

In dieser Arbeit konnte gezeigt werden, dass eine multimodale (Audio- und Bewegungsdaten) Echtzeiterkennung von Lachen mit Smartphones möglich ist und hierfür auch Daten, die nicht unter Laborbedingungen aufgezeichnet wurden, ausreichend sind. Praktisch wurde das mobileSSI-Framework, mit Erweiterungen für die Android-Plattform, eingesetzt und auch dessen verteilte Nutzung über mehrere Geräte demonstriert. Da alle Daten auf dem Gerät des Nutzers aufgezeichnet und verarbeitet werden können, wird dessen Privatsphäre geschützt.

6.3.3 Biofeedback des Herzschlags durch Tangible-Displays zur Verstärkung des eigenen und geteilten affektiven Erlebnisses

Einleitung

In diesem Abschnitt wird eine Untersuchung vorgestellt, bei der ein eigenes Biofeedback-System entwickelt und die Auswirkungen des Systems auf das Nutzererlebnis evaluiert wurden. Generell wäre dieses System auch für Meditationsübungen zur Stressreduktion in Gruppen geeignet, wodurch es negative Effekte auf das mentale und emotionale Wohlbefinden abbauen und zugleich positive Effekte auf das soziale Wohlbefinden ausüben könnte.

In dieser Arbeit wurde ein „Tangible-Display“ in ähnlicher Form und taktilem Feedback eines menschlichen Herzens entwickelt und für Biofeedback eingesetzt. Hierdurch konnte der Herzschlag einer Person durch das in Händen gehaltene künstliche Herz fühlbar gemacht werden. Der von einer Person erfasste Puls konnte hierbei derselben Person ausgegeben oder an das künstliche Herz einer weiteren Person übertragen werden. Um zu untersuchen, wie ein derartiges Tangible-Display das Nutzererlebnis im Alltag beeinflussen kann, wurde eine Feldstudie mit 60 Nutzern (30 Paare) durchgeführt, bei der kurze Filme gezeigt wurden. In jeder Sitzung wurde zwei Nutzern jeweils ein Prototyp gegeben und drei Konditionen getestet: 1) das Tangible-Display zeigte den eigenen Herzschlag, 2) es zeigte den Herzschlag des Partners und 3) es wurde kein Tangible-Display eingesetzt. Bei der Studie konnten statistisch signifikante Effekte bei der Empathie festgestellt werden, wenn der Herzschlag des Partners angezeigt wurde. Signifikante Effekte traten ebenfalls bei der Immersion, wenn der eigene Herzschlag abgebildet wurde, auf.

Diese Untersuchung wurde in den Publikationen (Aslan, Seiderer, Dang, Rädler & André, 2020a, 2020b) veröffentlicht.

Hintergrund

Achtsamkeit und psychologisches Wohlbefinden Viele Aktivitäten der inneren Organe sind für den Menschen nicht bewusst wahrnehmbar. Bei gesunden Personen gibt es gute Gründe, diese Details auszublenden, um sich von diesen nicht ablenken zu lassen. Allerdings kann ein bewusstes Reflektieren über diese inneren Abläufe auch dabei helfen, ungesunde Verhaltensmuster zu erkennen und einzudämmen. Viele traditionelle Achtsamkeitspraktiken stehen Technik eher skeptisch gegenüber, da diese ablenkend wirken kann. Gleichzeitig gibt es in der Informatik allerdings verstärkt Entwicklungen, die zur Verbesserung des Wohlbefindens, der mentalen Gesundheit und Erhöhung von Achtsamkeit beitragen können. Ein Beispiel ist in der Arbeit von Gervais et al. (2016) zu finden, in der mittels eines anthropomorphen, physischen Avatars physiologische (z. B. Herzrate) und davon abgeleitete Daten (z. B. Arbeitsbelastung) visuell dargestellt wurden. Roo et al. (2017) präsentierten einen augmentierten Sandkasten, der den inneren Zustand eines Nutzers abbildete und zeigte, wie sich dieser verändert. Nutzer konnten den Sand mit ihren Händen verändern. Taktiler Feedback erfolgte allerdings nicht. Für Atemübungen wurden virtuelle Avatare und Spiele z. B. von Sonne und Jensen (2016) eingesetzt.

Meditationsübungen zur Stressreduktion Ein kosteneffektives, etabliertes Programm zur Stressreduktion stellt die achtsamkeitsbasierte Stressreduktion (engl. Mindfulness-Based Stress Reduction = MBSR) dar, das von Kabat-Zinn (1982) entwickelt wurde. Das Programm kann laut Kabat-Zinn (1982) auch in Gruppen angewandt werden, wobei hierdurch die Effizienz verbessert wird, da sich die Teilnehmer gegenseitig helfen können, indem sie ihre Erfahrungen teilen und sich motivieren. Laut Shapiro et al. (2005) basiert MBSR auf der Meditationsform der Achtsamkeit (engl. mindfulness), die ursprünglich von einer 2500 Jahre alten buddhistischen Tradition abgeleitet wurde. Laut den Autoren lernen Teilnehmer des Programms sich Gedanken, Gefühlen und Körperempfindungen bewusster zu werden. Durch die MBSR-Intervention können Stress, Angst und Depression reduziert (Shapiro et al., 2005) und das Wohlbefinden verbessert (Nyklíček & Kuijpers, 2008) werden. Technologie kann bei solchen meditativen Übungen unterstützend eingesetzt werden, um beispielsweise den eigenen Herzschlag leichter spürbar zu machen (Biofeedback). Laut Slovák et al. (2012) kann durch das Teilen der Herzrate mit einer bekannten Person die gefühlte soziale Verbundenheit zwischen den Personen erhöht werden, da das Teilen dieser Information als eine Art emotionale Selbstenthüllung angesehen werden kann. Soziale Verbundenheit besitzt ebenfalls positive Effekte auf die Gesundheit sowie das Wohlbefinden (siehe Abschnitt 6.1.3).

Tangible Interfaces für die Herzrate und Nutzererlebnis in sozialem Kontext In der Literatur sind diverse Arbeiten zu finden, in denen Tangible-Displays zur Darstellung der Herzrate eingesetzt

wurden. Am häufigsten wurde zur Bestimmung der Herzrate ein Armband mit einem optischen Sensor genutzt. Ein Beispiel hierfür ist die Arbeit von Costa et al. (2016), welche die Entwicklung eines tragbaren Geräts mit dem Namen „EmotionCheck“ beschreibt. Mit dem Gerät wurde die Herzrate am Handgelenk bestimmt sowie ein Vibrationsmuster an der Unterseite des Geräts erzeugt. Es wurde hierbei festgestellt, dass ein verfälschtes Feedback sogar die Emotionen eines Nutzers beeinflussen können.

In der Literatur sind verschiedene Tangible Interfaces aufzufinden, bei denen die Herzrate eingesetzt wurde, um das Nutzererlebnis zu beeinflussen. Diese Arbeiten können in folgende Kategorien unterteilt werden: Spiele, Kunst, Sensibilisierung (engl. awareness) und (örtlich entferntes) geteiltes Erlebnis.

Ein Beispiel für ein sogenanntes „Serious Game“ wurde in der Arbeit von Al Osman et al. (2016) vorgestellt, bei dem ein Spiel zur Stressreduktion mittels des Herzschlags gesteuert wurde. In der Literatur finden sich noch viele weitere Beispiele, die zum Teil auch kompetitiv sind. Bei der Entwicklung des eigenen Systems stand allerdings ein geteiltes Nutzererlebnis in einer kollaborativen oder kooperativen Weise im Vordergrund.

Núñez-Pacheco und Loke (2014) stellten eine künstlerische Arbeit vor, bei der die Herzrate einer Person auf deren Kleid mittels farbiger Kreise projiziert wurde. Es konnte hierbei festgestellt werden, dass Nutzer, die bewusst versuchten, ihre Herzrate zu verändern, eine stärkere Verbundenheit mit dem Erlebnis empfanden als Zuschauer. Diese Veröffentlichung verfolgte ein ähnliches Ziel wie das in der eigenen Arbeit vorgestellte System, welches das Bewusstsein über die eigene Herzrate stärken sollte. Hervorzuheben ist, dass die eigene Arbeit nicht ausschließlich für ein individuelles, sondern für ein geteiltes Erlebnis zwischen zwei Nutzern konzipiert wurde, um Auswirkungen auf deren soziale Verbundenheit zu untersuchen.

Geteiltes und / oder örtlich entferntes Erlebnis wurde in der Veröffentlichung von Walmink et al. (2014) behandelt. Sie verwendeten Anzeigen, die auf der Rückseite von Fahrradhelmen angebracht waren, um die aktuelle Herzrate des entsprechenden Fahrradfahrers darzustellen. Auf diese Weise kannten alle Fahrradfahrer einer Gruppe die Herzrate aller Gruppenmitglieder, was ein geteiltes Erlebnis ermöglichte und das Engagement verstärkte. In der Arbeit von J. Werner et al. (2008) wurde ein Prototyp vorgestellt, bei dem der Puls mittels zwei Ringen zwischen zwei Personen an verschiedenen Orten geteilt werden konnte. Bereits eine minimale taktile Ausgabe war ausreichend, um eine Sinneserfahrung auszulösen. Slovák et al. (2012) untersuchten in einer Studie, wie das Feedback der eigenen Herzrate oder von örtlich entfernten Personen verstanden wurde. Die Arbeit legt nahe, dass eine Kommunikation per Herzrate die soziale Verbundenheit oder andere Arten sozialer Interaktion unterstützen kann.

Es existieren weitere Arbeiten, bei denen die Herzrate visuell oder auditiv ausgegeben wurde. In der vorliegenden Arbeit wurde allerdings primär die taktile Modalität untersucht.

Woodward et al. (2018) präsentierten einen Prototyp, der ein geteiltes Erlebnis mittels taktilem Feedback durch einen Würfel und ein Armband ermöglichen sollte. Es wurde in diesem Fall jedoch, im Gegensatz zur eigenen Arbeit, nicht die Herzrate übermittelt, sondern vier verschiedene Arten von Emotionen erkannt und in Vibrationsmuster übersetzt.

Der in dieser eigenen Arbeit entwickelte Prototyp sollte die Herzrate auf eine möglichst intuitive Art ausgeben. Hierbei eignete sich eine greifbare und fühlbare Nachbildung des Herzens, die den eigenen Herzschlag spürbar machen kann. Prototypen in dieser physikalischen Form wurden bisher in der Literatur nicht präsentiert. Die hier vorgestellten verwandten Arbeiten zeigen allerdings, dass Tangible-Displays geeignet sind, Emotionen, das individuelle Nutzererlebnis und Selbstbewusstsein zu beeinflussen. Zudem konnten durch das Teilen der Herzrate Auswirkungen auf das geteilte Erlebnis und die soziale Verbundenheit festgestellt werden. Ziel des eigenen Prototyps war es, dessen Auswirkungen auf das eigene und geteilte Nutzererlebnis zu untersuchen, was in anderen Arbeiten nicht in einer Laborstudie kombiniert analysiert wurde. Die Studienergebnisse können hierbei bedeutsame Hinweise für zukünftige Anwendungen zur Verbesserung des Wohlbefindens durch Stressreduktion oder Erhöhung der sozialen Verbundenheit liefern.

Design und Entwicklung des Prototyps

Der Entwicklungsablauf des Designs war vergleichbar mit dem von Loke und Robertson (2013) vorgeschlagenen verkörpertten Ansatz für bewegungsbasiertes Interaktionsdesign. Für den Designprozess wurde eine explorative Herangehensweise gewählt, wie sie in der Designforschung vorgeschlagen wird (Zimmerman et al., 2007). Daher wurde vergleichsweise viel Zeit investiert, um verschiedene Designs des künstlichen Herzens zu testen, wie es sich in der Hand gehalten anfühlt und welches Design am besten geeignet ist, die eigene Herzrate spürbar zu machen.

Das Design des künstlichen Herzens sollte dessen Gewicht, Form, Dimensionen, Fragilität und Bewegungen widerspiegeln. Das künstliche Herz basierte auf der Arbeit von Aslan et al. (2016). Sie lieferte Erkenntnisse, auf die bei der Herstellung des neuen Herzprototyps zurückgegriffen werden konnte. Die innere mechanische Konstruktion aus zwei PET-Plastikringen und einem Servomotor wurde übernommen. Auf die Nutzung von Latex für die Umhüllung wurde allerdings verzichtet, da sie einen intensiven Geruch verursachte. Beim neuen Prototyp wurde daher Silikon verwendet. Um eine relativ realistische Außenhülle mit Adern zu erhalten, wurde das Silikon mithilfe einer, mit einem 3D-Drucker gedruckten, Negativform gegossen. Es wurden verschiedene Härtegrade an Silikon mittels Proben getestet und schlussendlich ein sehr flexibles, lebensmittelechtes Silikon ausgewählt. Die Farbe des Silikons war grün, da eine realistische herzförmige Farbe möglicherweise nicht gewollte Gefühle auslösen könnte. Die Gussform wurde mit ausreichend flüssigem Silikon teilweise befüllt und für mehrere Minuten in alle Richtungen rotiert bis es ausgehärtet war, um eine möglichst gleichmäßig dicke Hülle zu erhalten, die von der Mechanik (Plastikringe und Servomotor) noch ausreichend gut verformt werden konnte.

Die Elektronik des Prototyps wurde im Vergleich zu (Aslan et al., 2016) ebenfalls vollständig neu entwickelt. Da das System nun zwei Raspberry Pi Zero W nutzte, wurde es insgesamt als „PiHeart“-System bezeichnet.

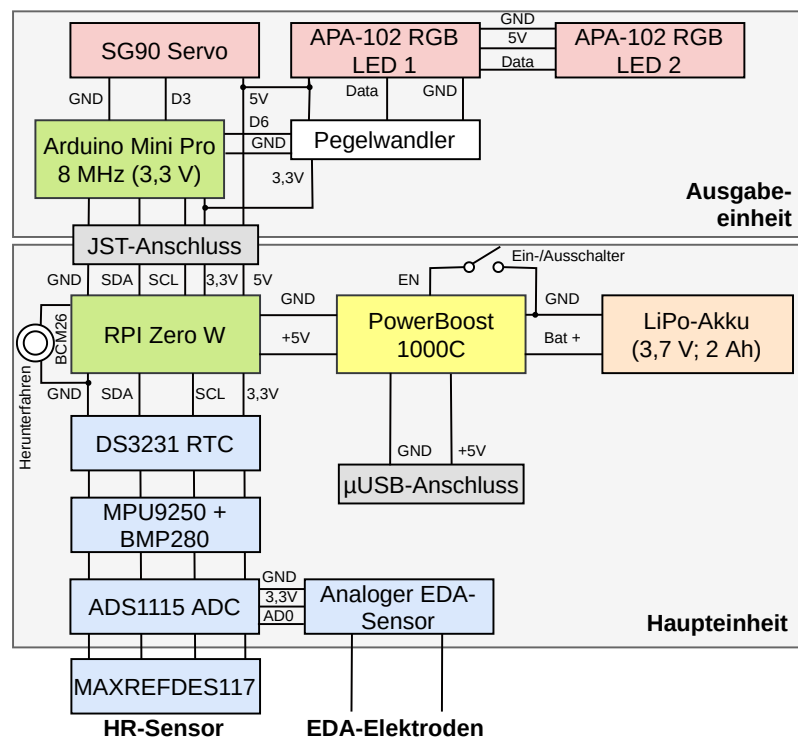


ABBILDUNG 6.7: Schaltplan der erweiterten Version des PiHeart.

In Abbildung 6.7 ist der vereinfachte Schaltplan für den PiHeart-Prototyp dargestellt. Der Prototyp bestand aus zwei Komponenten, einer mobilen Haupt- und einer Ausgabeeinheit. Die

Haupteinheit enthielt einen Raspberry Pi Zero W, Sensorik und die Stromversorgung für alle Komponenten. Die Ausgabeeinheit beinhaltete einen Arduino Mikrocontroller, der zwei RGB-LEDs sowie einen Servomotor ansteuerte. Die RGB-LEDs wurden integriert, da die vorherige Version der Herzen zusätzlich eine visuelle Ausgabe besaß. Für diese Studie wurden sie allerdings per Software deaktiviert. Beide Einheiten waren mittels einer Steckverbindung miteinander verbunden.

Haupteinheit Abbildung 6.8 zeigt Bilder der vollständigen Haupteinheit aus verschiedenen Perspektiven.

Die Haupteinheit verfügte über einen LiPo-Akku mit 2 Ah, der an die Adafruit PowerBoost-1000C-Schaltung⁴ angeschlossen war. Diese Schaltung enthielt einen TPS61090 Boost Converter und war fähig, den Akku auch bei Nutzung gleichzeitig über USB zu laden. Die Schaltung lieferte eine Spannung von 5 V bei 2 A Stromstärke. Die Stromversorgung konnte mittels eines Schalters ein- oder ausgeschaltet werden. Der Gesamtstromverbrauch inklusive Ausgabeeinheit konnte mittels Datenblättern der Einzelkomponenten abgeschätzt werden und betrug maximal 1 A. Die Stromversorgung war daher bei vollgeladenem Akku für mindestens zwei Stunden gewährleistet.

An den Raspberry-Pi-Zero-W-Einplatinencomputer waren alle anderen Komponenten angeschlossen und wurden teilweise auch über diesen mit einer Spannung von 3,3 V versorgt. Er erfasste die Sensordaten und steuerte alle Komponenten.

An den Raspberry waren in der erweiterten Version des PiHeart folgende Komponenten per I²C angeschlossen: eine DS3231-Echtzeituhr, ein MPU9250-Multisensor bestehend aus je einem Accelerometer / Gyroskop / Magnetometer mit je 3-Achsen, ein BMP280 Luftdruck- / Temperatursensor, ein ADS1115-16-bit-Analog-Digital-Wandler und ein MAXREFDES117, das einen MAX30102-Herzratensensor enthielt. An den Analog-Digital-Wandler war eine analoge Schaltung zur Erfassung des Hautleitwerts angeschlossen. Bei der vereinfachten Variante der Schaltung fehlten folgende Komponenten: MPU9250, BMP280, ADS1115 sowie die analoge EDA-Schaltung. Die zusätzlichen Sensoren der erweiterten PiHeart-Version wurden in den Untersuchungen nicht eingesetzt und waren daher für mögliche weiterführende Studien vorgesehen.

Mittels einer JST-Steckverbindung wurde eine I²C-Schnittstelle für externe Komponenten herausgeführt. Zur externen Stromversorgung wurden zusätzlich 5 V und 3,3 V bereitgestellt, wobei die 5 V direkt vom Boost-Konverter stammten und eine höhere Stromstärke geliefert werden konnte.

Ausgabeeinheit In Abbildung 6.9 ist ein Bild der Ausgabeeinheit ohne die Herzsilikonhülle, die Mechanik und das Gehäuse dargestellt.

Die Ausgabeeinheit konnte optional, mittels einer fünfpoligen JST-Steckverbindung, an die Haupteinheit angeschlossen werden. Es hätten bei Bedarf auch andere Komponenten, die den I²C-Bus verwenden, verbunden werden können. Hierdurch hätten beispielsweise andere Herzprototypen problemlos getestet werden können.

In diesem Fall bestand die Ausgabeeinheit aus einem Arduino Mini Pro mit 8 MHz, der mit 3,3 V betrieben wurde, einem SG90-Servomotor sowie zwei APA-102 RGB-LEDs, die über einen Signalpegelwandler mit dem Arduino verbunden waren. Der vereinfachten Version der Ausgabeeinheit fehlten die beiden RGB-LEDs. Der Arduino war so programmiert, dass dieser als I²C-Slave fungierte, der externe Befehle, hier vom Raspberry, empfing und ausführte.

Der Servomotor wurde direkt über die 5 V Spannung mit Strom versorgt, da diese mehr Leistung abgeben konnte. Das PWM-Steuersignal stammte direkt vom Arduino und konnte hierbei problemlos bei 3,3 V liegen. Bei den RGB-LEDs (TTL-Logik = 5 V) war ein direkter Anschluss nicht möglich, daher war ein Signalpegelwandler vorgeschaltet, der das Datensignal von 3,3 V auf 5 V wandelte. Die zweite RGB-LED war mit dem Ausgabe-Datenkanal der ersten LED verbunden. Beide LEDs waren über einen Bus einzeln ansteuerbar.

⁴<https://learn.adafruit.com/adafruit-powerboost-1000c-load-share-usb-charge-boost/overview> (12.03.2020)

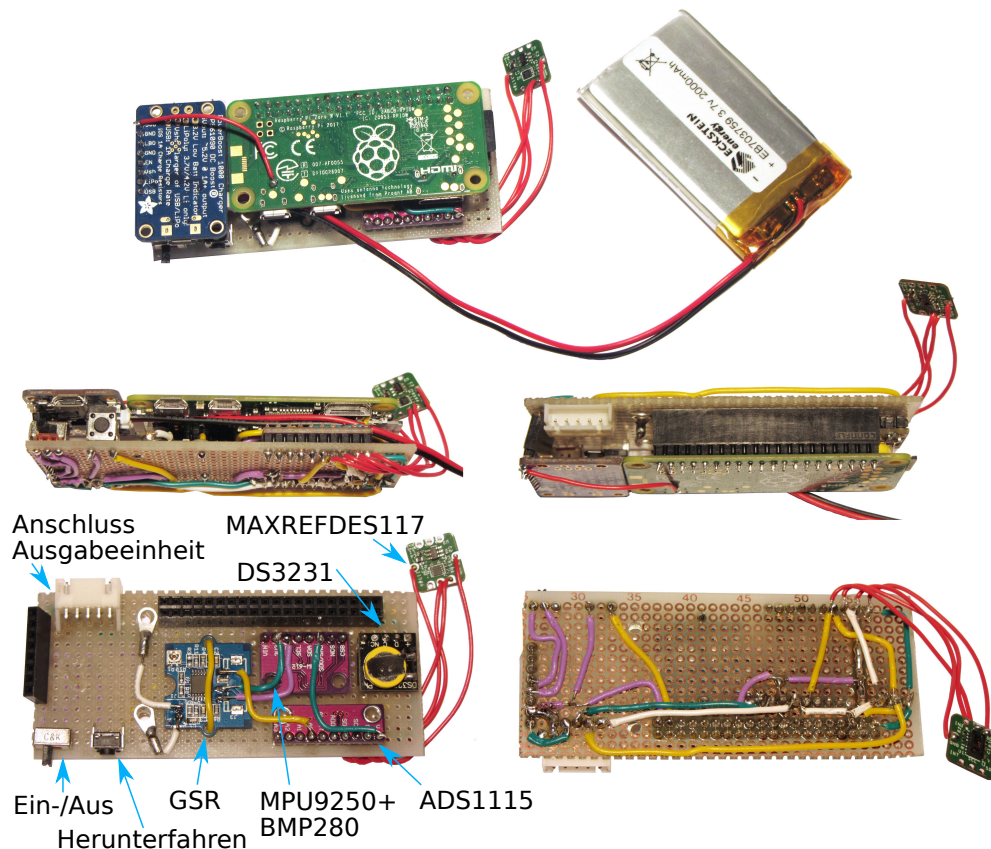


ABBILDUNG 6.8: Platine der Haupteinheit aus verschiedenen Perspektiven. Für die unteren beiden Bilder wurden der Raspberry und die PowerBoost 1000C Platine abgezogen.

Gehäuse und Hülle Zum Schutz der elektronischen Komponenten und für eine sichere Verwendung (Lithium-Akku) mussten stabile Gehäuse für die Prototypen entwickelt werden. Um möglichst passgenaue Gehäuse zu erhalten, wurden für diese zunächst maßhaltige 3D-Modelle erstellt, gesliced und anschließend mit einem Ultimaker-2+-3D-Drucker⁵ ausgedruckt. In Abbildung 6.10 links ist das Gehäuse der Haupteinheit des PiHeart-Prototyps mit seinen Aussparungen dargestellt. Um Platz zu sparen, wurde das Gehäuse mittels im geschlossenen Zustand durchgefädelter Klettbänder zusammengehalten. Das Gehäuse für die Komponenten der Ausgabeeinheit, die nicht im Herz enthalten waren, wurde auf dieselbe Art wie das Gehäuse der Haupteinheit erstellt. Dieses war jedoch mit Blechschrauben verschraubt. Abbildung 6.10 rechts zeigt alle fertigen Komponenten eines PiHeart-Prototyps mit den zugehörigen Gehäusen.

Software

Für den Raspberry-Einplatinencomputer und AVR-Mikrocontroller wurden vom Autor dieser Arbeit spezielle Software entwickelt und implementiert, die in den folgenden Abschnitten genauer beschrieben wird.

Arduino Für den Arduino der Ausgabeeinheit wurde eine Firmware entwickelt, die es erlaubte diesen als I²C-Slave am I²C-Bus des Rasperrys zu nutzen. In der Firmware wurden folgende externe Bibliotheken genutzt: *VarSpeedServo*⁶ für den Servo und *FastLED*⁷ für die RGB-LEDs.

⁵<https://ultimaker.com/en/products/ultimaker-2-plus> (12.03.2020)

⁶<https://github.com/netlabtoolkit/VarSpeedServo.git> (12.03.2020)

⁷<http://fastled.io/> (12.03.2020)

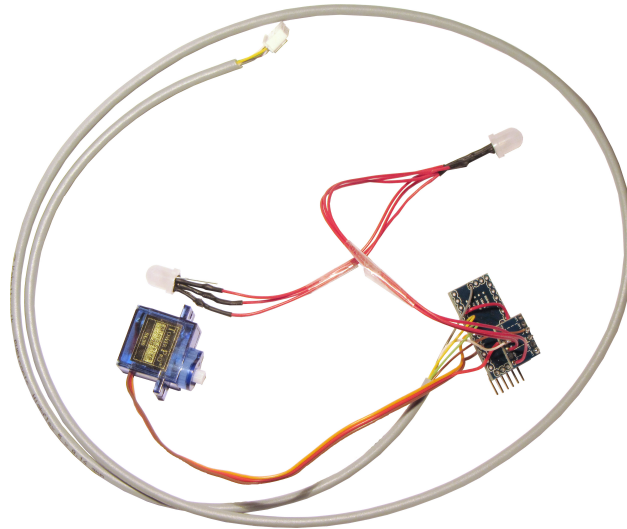


ABBILDUNG 6.9: Ausgabeeinheit des Prototyps ohne Silikonhülle und Gehäuse. Der Arduino Mini Pro mit Pegelwandler sowie die beiden RGB-LEDs und der Servomotor sind erkennbar.

Die Firmware des Arduino besaß nur einen rudimentären Funktionsumfang und wurde von komplexerer Software auf dem Raspberry angesteuert.

Der Arduino hatte hierfür eine eigene I²C-Adresse erhalten, worüber er, wie alle anderen Geräte angesteuert werden konnte. Sobald über die I²C-Schnittstelle ein Event empfangen wurde, erfolgte ein Herzschlag, der realisiert wurde, indem sich die Achse des Servos von Position 0 auf 180 und wieder zurück auf 0 drehte. Der Servo bewegte sich hierbei mit seiner Maximalgeschwindigkeit.

Raspberry Pi Zero W Das offizielle Raspbian-Linux wurde als Betriebssystem für die Raspberry Pis genutzt. Beim Starten wurde automatisch eine WiFi-Verbindung zu einer vorgegebenen SSID aufgebaut. Da ein Avahi-Service⁸ auf dem System aktiv war, konnte der entsprechende Raspberry über seinen Rechnernamen direkt im Netzwerk gefunden werden.

Die eigenen Programme wurden mit Python 3 entwickelt und kommunizierten im Netzwerk mit anderen Programmen per MQTT-Protokoll. Auf jedem Raspberry wurde hierfür eine lokale Instanz des MQTT-Brokers *Mosquitto*⁹ ausgeführt.

Ansteuerung Ausgabeeinheit Für die Ansteuerung der Ausgabeeinheit wurde ein Programm entwickelt, das per I²C-Bus mit einer bestimmten Herzrate den Befehl zum Ausführen eines Herzschlages an den Arduino sendete. Die Herzrate wurde hierbei per MQTT-Protokoll empfangen, woraufhin ein Timer mit dem entsprechenden Intervall gestartet wurde. Der Timer konnte jederzeit gestoppt werden, indem eine Herzrate von 0 bpm gesendet wurde.

Bestimmung der Herzrate

Um die Herzrate des Nutzers messen zu können, wurde ein Programm entwickelt, das zunächst den Herzratensensor (MAX30102) aktivierte und konfigurierte. Es erfasste zudem die BVP-Daten per I²C und berechnete die Herzrate mittels Fouriertransformation (Details siehe Abschnitt 3.2.2). Die Herzrate wurde abschließend per MQTT-Protokoll versendet.

⁸<https://www.avahi.org/> (12.03.2020)

⁹<https://mosquitto.org/> (12.03.2020)

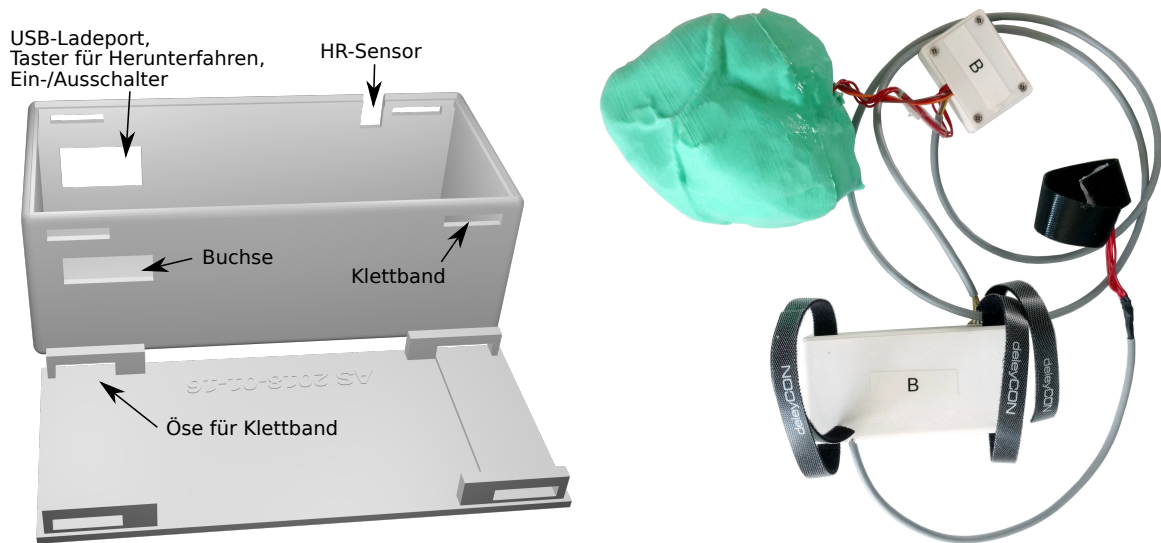


ABBILDUNG 6.10: Links: 3D-Ansicht des Gehäuses mit beschrifteten Aussparungen. Rechts: Alle Komponenten eines PiHeart-Prototyps.

Feldstudie

Um zu untersuchen, wie die Verwendung der PiHearts das Nutzererlebnis in einem alltäglichen sozialen Umfeld beeinflussen, wurde eine Feldstudie mit Teilnehmerpaaren durchgeführt. Da das gemeinsame Ansehen von Filmen ein alltägliches Erlebnis darstellt, wurde dieses Szenario gewählt. Die Studie sollte auf folgende Forschungsfrage antworten geben: Wie verändert die Nutzung der PiHearts das Nutzererlebnis, während gemeinsam ein Film angesehen wird, wenn die eigene Herzrate oder die des Partners ausgegeben wird? Die Erwartung vor der Studie war, dass die PiHearts die Beziehung zum eigenen Körper und die soziale Bindung stärken können.

An der Studie nahmen 30 Teilnehmerpaare (60 Personen) teil (30 weiblich, 30 männlich). 24 Paare gaben an, in einer Partnerschaft zu sein, 4 Paare waren Freunde und zwei Paare waren Familienmitglieder (z. B. Schwester oder Bruder). Die Altersspanne der Teilnehmer war relativ groß, beispielsweise waren 14 Personen zwischen 45 und 65 Jahre alt und 17 Teilnehmer zwischen 18 und 24. Das mittlere Alter betrug ca. 31,4 Jahre mit einer Standardabweichung von ca. 13,7 Jahren.

Neben den PiHearts wurde bei der Studie ein Laptop eingesetzt. Der Aufbau ist in Abbildung 6.11 dargestellt. Es wurden drei kurze Filme ausgewählt, die einen vergleichbaren Aufbau besaßen und bei Bewertungsportalen wie „IMDb“ ähnliche Bewertungen erhielten: „Big Buck Bunny“, „Overwatch“ (Trailer) und „Der Vogelschreck“ (Trailer).

Aufbau Für die Studie kommunizierten die PiHeart-Prototypen per MQTT mit einer Software in C# auf einem Windows-Laptop. Diese steuerte die PiHearts, zeichnete die Daten auf und zeigte die Videos auf dem Laptopdisplay an. Abbildung 6.12 gibt einen Überblick über die Kommunikation zwischen den Geräten und der Software. Das auf dem Laptop ausgeführte Programm speicherte in einer SQLite-Datenbank die Herzrate, den angezeigten Filmtitel, die Modalität und die BVP-Daten mit Zeitstempeln, mit denen die Daten auch synchronisiert wurden. Die Software konnte per SSH-Verbindung bei Beginn einer Studien-Sitzung die notwendigen Skripte auf den Raspberry Pis starten und danach per MQTT ansteuern. Die MQTT-Broker wurden jeweils auf den Raspberry Pis ausgeführt, sodass kein zentraler Server mit einem bestimmten Namen auf einem Laptop / PC im Netzwerk benötigt wurde. Ein kompakter WiFi-Router mit einer Stromversorgung per USB wurde zur kabellosen Kommunikation zwischen den Geräten eingesetzt.



ABBILDUNG 6.11: Studienaufbau in der Wohnung der Studienteilnehmer (Bild: Simon Rädler).

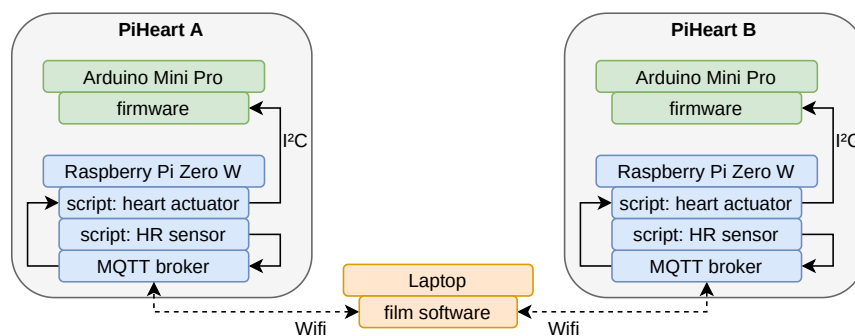


ABBILDUNG 6.12: Kommunikation zwischen den Geräten und Programmen.

Ablauf Zunächst wurde eine Pilotstudie mit zwei Teilnehmern durchgeführt, um ggf. Fehler im Ablauf und der Software zu entfernen. Daraufhin erfolgte die Studie innerhalb von vier Wochen mit den 30 Teilnehmerpaaren. Es wurden drei Konditionen mithilfe einer „within-subject“-Studie untersucht. Mittels eines Fragebogens zum Nutzererlebnis wurden von jedem Teilnehmer für jede Kondition Daten erfasst. Der Fragebogen, der für die Messung des Spielerlebnisses entwickelt wurde, stammte aus der Publikation von IJsselsteijn et al. (2013). Er beinhaltet verschiedene Fragen, die relevant sind, um das Erlebnis kontinuierlicher Unterhaltungsaktivitäten zu beurteilen. Zudem behandelt der Fragebogen soziale Aspekte, die in der Studie untersucht werden sollten. Es wurden die folgenden drei Konditionen während des Ansehens eines Filmes durchgewechselt: 1) das PiHeart-System zeigte jeweils den eigenen Herzschlag, 2) es zeigte den Herzschlag des Partners und 3) es wurde kein Tangible-Display eingesetzt. Die Reihenfolge der Konditionen war ausgeglichen (counter balanced), während die Filmreihenfolge identisch war. Für die Studie wurde nicht nur ein Film ausgewählt, da das wiederholte Ansehen desselben Kurzfilmes einen starken, vermutlich eher negativen Einfluss auf das Nutzererlebnis gehabt hätte. Am Ende der Studie wurde ein semistrukturiertes Interview durchgeführt. Die Auswertung der Antworten bei den Befragungen wurden mittels thematischer Analyse (Braun & Clarke, 2006) durchgeführt, bei denen die für die Forschungsfragen relevanten Themen identifiziert und nach der Häufigkeit der Antworten sortiert wurden.

Ergebnisse Die Abbildungen 6.13 und 6.14 zeigen die **Ergebnisse des Fragebogens**, wobei Abbildung 6.13 die Bewertungen der Studienteilnehmer bezüglich ihres Erlebnisses während des Films darstellt und Abbildung 6.14 deren Gefühle hinsichtlich der sozialen Präsenz. Tabelle 6.5 zeigt die Ergebnisse einer wiederholten Varianzanalyse (ANOVA). Im Fall von signifikanten Ergebnissen wurden die Resultate paarweiser Post-Hoc-Tests (Bonferroni-Korrektur) angegeben.

Es zeigte sich ein Haupteffekt (engl. main effect) beim Messwert „sensorische und imaginative Immersion“. Die paarweisen Vergleiche ergaben, dass es ein signifikanter Unterschied zwischen den Konditionen „mit eigenem Herz“ und „ohne Herz“ auftrat. Das bedeutet, dass die Teilnehmer eine signifikant höhere Immersion wahrnahmen, wenn sie ein PiHeart mit eigenen Herzschlägen in der Hand hielten, als ohne künstliches Herz. Bezüglich der sozialen Präsenz konnte ein Haupteffekt bei allen drei Messwerten („Empathie“, „verhaltensbezogene Beteiligung“ und „negative Gefühle“) festgestellt werden. Die paarweisen Vergleiche zeigten, dass es einen signifikanten Unterschied zwischen der Kondition, wenn Studienteilnehmer ein PiHeart mit dem Herzschlag des Nachbarn / Partners hielten („mit Nachbar-Herz“) und den anderen beiden Konditionen gab. Wenn die Teilnehmer die Filme mit dem Herzschlag des Nachbarn ansahen, fühlten sie ein signifikant höheres Niveau sozialer Präsenz.

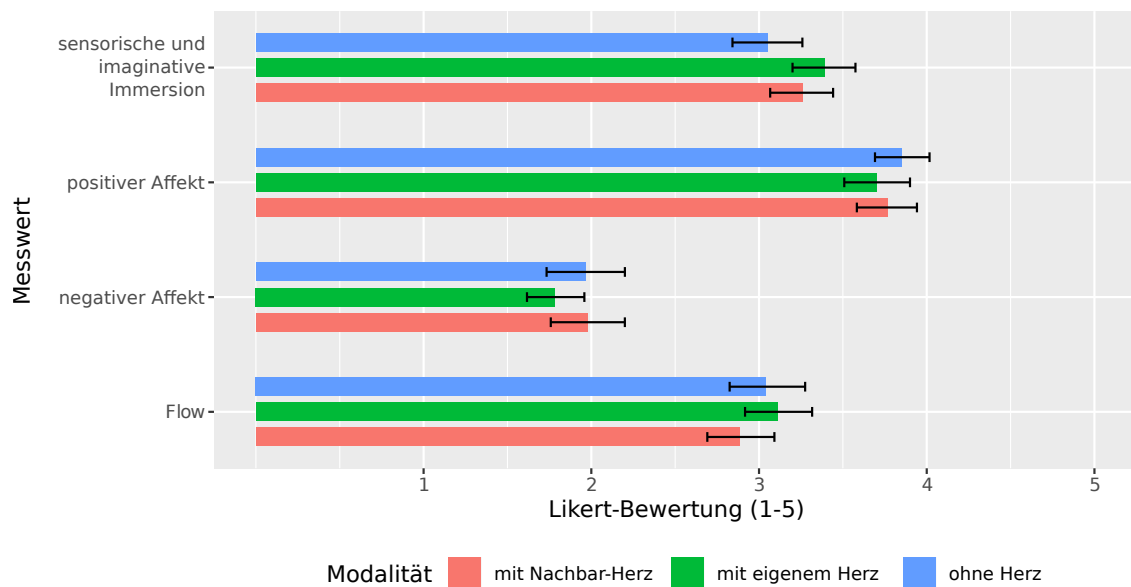


ABBILDUNG 6.13: Ergebnisse des Fragebogens.

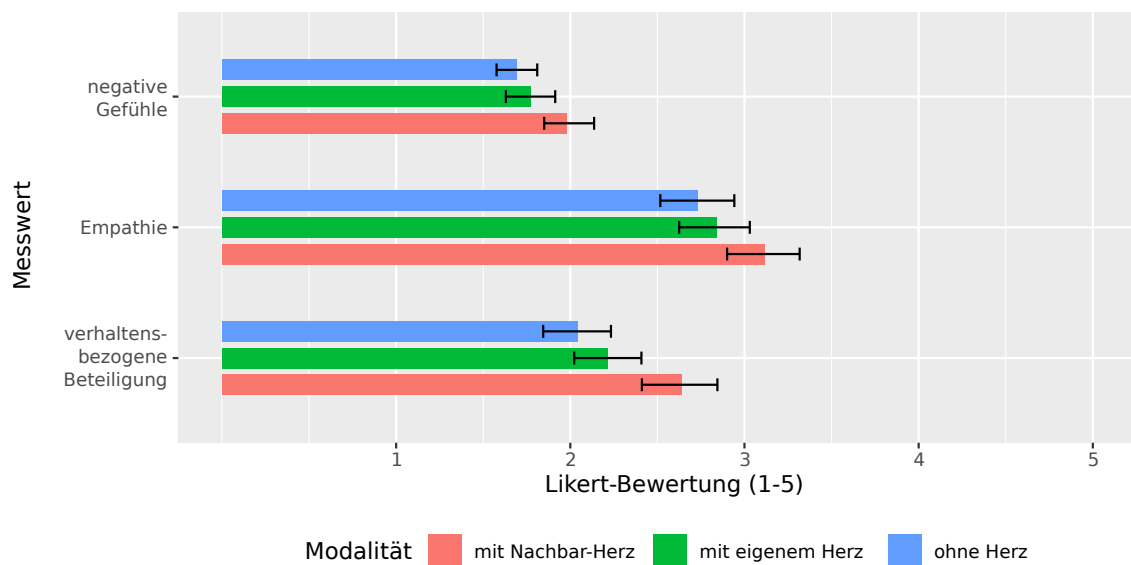


ABBILDUNG 6.14: Ergebnisse bezüglich sozialer Aspekte des Fragebogens.

Zusätzlich wurden **semistrukturierte Interviews** mit den Studienteilnehmern am Ende jeder Sitzung durchgeführt, bei denen sie offene Fragen beantworteten, wie z. B. was den Probanden am

abhängige Variable	F-Wert	p-Wert	paarweise Post-Hoc-Tests (Bonferroni-Korrektur)
positiver Affekt	F=1,40	p=0,25	
negativer Affekt	F=1,69	p=,19	
sensorische und imaginative Immersion	F=4,78	p=,01 *	O-N (p=,32); E-N (p=,63); E-O (p=,003)
Flow	F=2,06	p=,13	
Empathie	F=8,55	p<,001 ***	O-N (p<,001 ***); E-N (p=,02 *); E-O (p=,92)
verhaltensbezogene Beteiligung	F=16,71	p<,001 ***	O-N (p<,001 ***); E-N (p=,0015 **); E-O (p=,29)
negative Gefühle	F=7,53	p<,001 ***	O-N (p<,001 ***); E-N (p=,03 *); E-O (p=,96)

TABELLE 6.5: Überblick über die statistischen Tests über die Daten aller Studienteilnehmer, wobei der Gesamt-Effekt der verschiedenen Konditionen auf verschiedene Messwerte des Fragebogens angegeben sind. Zusätzlich sind die paarweisen Vergleiche mittels des Post-Hoc-Tests enthalten. Abkürzungen für Signifikanzen: *** = 0,001; ** = 0,01; * = 0,05. Abkürzungen für paarweise Vergleiche: ohne Herz - mit Nachbar-Herz (O-N), mit eigenem Herz - mit Nachbar-Herz (E-N), mit eigenem Herz - ohne Herz (E-O).

besten oder wenigsten gefiel. Die meisten Teilnehmer (Anzahl: 17) äußerten positives Feedback bezüglich der gewählten, spürbaren Ausgabemodalität für ihre Herzrate, die sie gegenüber der häufig genutzten alphanumerischen Darstellung präferierten. 14 Probanden fanden es gut, dass sie die Herzrate ihres Partners spüren konnten. Drei Teilnehmer erwähnten explizit eine Erhöhung der Selbstwahrnehmung bezüglich der Herzrate. Bei der Frage, was die Studienteilnehmer am wenigsten mochten, wurde von vielen einer der Filme genannt. Zwei Personen fühlten sich vom PiHeart abgelenkt. Am häufigsten (Anzahl: 13) berichteten die Probanden vom Problem, dass sie nervös wurden, wenn die Herzrate sich erhöhte.

Bezüglich möglicher Einsatzgebiete nannten 12 Studienteilnehmer gesundheitsbezogene Anwendungen und einen Einsatz zur Verbesserung der Körperwahrnehmung. Sieben Probanden konnten sich eine Nutzung bei der Paartherapie vorstellen.

Die Teilnehmer wurden auch gefragt, wie es sich für sie anfühlte, ihren eigenen Herzschlag zu spüren. Hierauf wurde beispielsweise bezüglich der Herzrate geantwortet, dass „es interessant war Änderungen zu spüren“ und, dass „es sich schneller anfühlte, als gedacht“. Bezüglich des gefühlten Herzschlags ihres Partners wurde interessanterweise am häufigsten geantwortet, dass „man sich so fühlt, als wäre man in einem Wettbewerb, bei dem man versucht die geringste Herzrate zu haben“. Dies erklärt, warum es ein signifikantes Ergebnis bei dem Messwert „negative Gefühle“ (häufig bezogen auf kompetitive Gefühle) gab, der Teil des Fragebogens (soziale Präsenz) war.

Diskussion Es zeigte sich in der Auswertung, dass die Studienteilnehmer im Vergleich zum Erlebnis ohne PiHeart eine signifikant höhere Immersion (Eintauchen) rückmeldeten, als wenn sie ihren eigenen Herzschlag durch ein PiHeart spüren konnten. Ein hohes Maß an Immersion ist das Ergebnis einer starken Beziehung zu Dingen oder Personen, mit denen man sich beschäftigt. Es könnte hierbei ein Zusammenhang darin bestehen, dass die eigene Herzrate, mit der des PiHearts übereinstimmte und somit die Selbstwahrnehmung erhöht wurde. Einen Einfluss darauf könnte auch die taktile Ausgabemodalität der Herzrate besessen haben, die von vielen Probanden in den Interviews positiv angesehen wurde.

Wenn der Herzschlag des Partners durch das PiHeart spürbar war, konnte ebenfalls ein signifikanter Effekt bezüglich der sozialen Präsenz mittels des Fragebogens festgestellt werden. Die Teilnehmer zeigten ein signifikant höheres Niveau bei 1) der Empathie (z. B. Verbundenheit zu ihrem Partner), 2) den negativen Gefühlen (z. B. Gefühle bei einem Wettbewerb) und 3) dem Mitwirken am Verhalten (z. B. Annahme, dass sich die Herzschläge gegenseitig anpassen).

1) / 2) Empathie wurde mit mehreren Fragen gemessen. Hierzu zählen: das Gefühl der Verbundenheit mit der anderen Person, sich glücklich fühlen, wenn die andere Person glücklich ist oder es erfreulich finden, mit der anderen Person zusammen zu sein. Der Messwert von „negativen

Gefühlen“ basierte auf Fragen bezüglich Eifersucht, Schadenfreude oder Rachsucht. Diese Gefühle schienen mit dem kompetitiven Verhalten der Studienteilnehmer zusammenzuhängen. Empathie und negative Gefühle sind Teil der verhaltensbezogenen Beteiligung. Letztendlich wurden beide Gefühle signifikant beeinflusst. Hierbei könnten Unterschiede in den Beziehungen zwischen den Nutzerpaaren eine Erklärung darstellen oder dass das PiHeart die Gefühle beider Personen verstärkt hat. Der gefühlte Herzschlag des Partners könnte hierbei eine Irritation oder eine Art Entfremdung verursacht haben, wenn dieser nicht synchron zum eigenen war. Dies wurde allerdings in dieser Studie nicht weiter untersucht und sollte zukünftig berücksichtigt werden, da ungewollte Entfremdung Beziehungen schädigen könnte.

3) Bei den Studienteilnehmern traten signifikant höhere Niveaus von verhaltensbezogener Beteiligung auf, wenn der Herzschlag des Partners gespürt wurde. Die verhaltensbezogene Beteiligung wurde durch Aussagen gemessen, inwieweit die eigenen Aktionen von denen der anderen Person abhängen und, ob eigene Aktionen die andere Person beeinflussten. Teilnehmer könnten erkannt haben, dass sie die Herzrate der anderen Person beeinflussen können oder spürten, dass sich die Herzrate der anderen Person an die eigene anpasst.

Zusammenfassung

Diese Untersuchung befasste sich mit der Wirkung von Biofeedback des Herzschlags hinsichtlich emotionalem, mentalem und sozialem Wohlbefinden. Hierfür wurde das Nutzererlebnis von Testpaaren bei der Betrachtung ausgewählter Kurzfilme erfasst und ausgewertet. Ein eigenentwickeltes künstliches Herz – in Händen gehalten – konnte als Tangible-Display den Herzschlag des Probanden oder (s)eines Partners fühlbar machen. Eine Feldstudie mit 60 Personen ergab, dass durch dieses Biofeedback eine intensivere Wahrnehmung des Films möglich ist und die Verbundenheit mit einem Partner gestärkt werden kann, also zur Verbesserung des sozialen Wohlbefindens und damit wahrscheinlich auch der Gesundheit beiträgt. Auch auf das mentale Wohlbefinden sind positive Auswirkungen durchaus möglich, da das bewusste Fühlen des eigenen Herzschlags z. B. im Rahmen einer Meditation eingesetzt werden könnte. In weiterführenden Untersuchungen könnte geklärt werden, ob das Verbundenheitsgefühl auch verstärkt wird, wenn die Partner räumlich getrennt sind oder die PiHearts von zwei fremden Personen genutzt werden. Letzteres wäre insbesondere von Interesse, da die Herzrate als intime Information betrachtet werden kann (Slovák et al., 2012).

6.4 Zusammenfassung des Kapitels

Dieses Kapitel beinhaltet Untersuchungen und hieraus resultierende Ergebnisse mittels Nutzung von IoT-Technologien und Mobile Computing zur Verbesserung des emotionalen, mentalen und sozialen Wohlbefindens.

Die Erfassung von Stress und dessen Auswirkungen, insbesondere auf das mentale Wohlbefinden, ist von großer Relevanz. Chronischer Stress kann gravierende, negative Auswirkungen auf das mentale Wohlbefinden besitzen. Hierzu zählen beispielsweise negative Emotionen wie Ärger, Angst oder Frustration. Es können allerdings auch langfristige Gesundheitsprobleme ausgelöst werden. Zur Erkennung von Stress wurde in dieser Arbeit eine multimodale Klassifizierung untersucht, wobei möglichst unaufdringliche Sensorik eingesetzt wurde, die in die Umgebung integriert werden kann.

Als Zielgruppe waren primär Büroarbeiter angedacht, da es im Büro durchaus häufig und schnell zu einer Überforderung kommen kann, wodurch Stress entsteht. Zur frühzeitigen Erkennung von Stress ist es hilfreich, IoT-Systeme und Mobile Computing einzusetzen, um z. B. rechtzeitig Pausen bzw. Meditationsübungen für den Stressabbau empfehlen zu können. Ein Büroarbeiter kann je nach Situation die anfallenden Aufträge in einem eher entspannten Zustand oder bei Zeitdruck unter Stressbelastung durchführen. Gewisse Aufträge erfordern auch höhere

geistige Konzentration, die jedoch nicht als Stress zu bewerten sind. Deshalb wurde in der eigenen Untersuchung im Vergleich zu anderen Arbeiten keine binäre Klassifikation durchgeführt, sondern die drei Klassen „entspannt“, „konzentriert“ und „gestresst“ unterschieden. Neben den etablierten physiologischen Daten (Blutvolumenpuls und Hautleitwert) als Referenz wurden in der Laborstudie auch Daten von **drucksensitiven Sitz- und Bodenmatten** sowie eines stationären Eyetrackers aufgezeichnet. Die Auswertungen zeigten, dass die Daten der verwendeten drucksensitiven Matten ausreichen können, um eine Stresserkennung zu ermöglichen. Insbesondere im Büro, wo in vielen Fällen an einem festen Arbeitsplatz gearbeitet wird, kann dies eine geeignete Alternative oder Ergänzung zur Erkennung mit Wearables darstellen. Die Druckmatten zeigen bezüglich des Schutzes der Privatsphäre deutliche Vorteile im Vergleich zu physiologischen-, Audio- oder Videodaten.

Einen bedeutsamen Indikator zur Ermittlung des emotionalen Wohlbefindens stellt die Erkennung positiver Emotionen dar. Eine positive Emotion kann hierbei durch Lachen ausgedrückt werden und ist ein starkes Signal bei der sozialen Interaktion, wobei eine Gefühlsansteckung stattfinden kann. Hierdurch wird die soziale Verbundenheit erhöht, was sich wiederum positiv auf das soziale Wohlbefinden auswirkt. Zusätzlich kann durch Lachen Stress abgebaut und somit das mentale Wohlbefinden ebenfalls positiv beeinflusst werden. In dieser Arbeit wurde eine **multi-modale, mobile Echtzeitklassifikation** von „Lachen“ und „nicht Lachen“ basierend auf Audio- und Beschleunigungsdaten (entscheidungsbasierte und ereignisbasierte Fusion) von Smartphones vorgestellt und evaluiert. Diese wurde mittels mobileSSI umgesetzt und war auf Smartphones (on device) ausführbar. MobileSSI konnte mit entsprechenden Plugins auch so verwendet werden, dass die Klassifikationsergebnisse mehrerer Nutzer mit Smartphones an ein Tablet per Netzwerk gesendet wurden, das diese grafisch darstellte. Die Untersuchungen zeigen die verschiedenen Möglichkeiten von (mobile)SSI zur Erfassung von Indikatoren für das Wohlbefinden auf. Durch die Nutzung der implementierten IoT-Protokolle wie MQTT kann zudem eine Vernetzung zwischen dem Mobile Computing und IoT-Bereich ermöglicht werden.

Schließlich wurde in diesem Kapitel das PiHeart-System vorgestellt, das aus eigener Software- / Hardware bestand und für taktilen Biofeedback der Herzfrequenz eingesetzt wurde. **Biofeedback** kann unter anderem für Meditationsübungen unterstützend angewandt werden, um Stress abzubauen und ermöglicht somit eine Verbesserung des mentalen Wohlbefindens. In der durchgeführten Feldstudie wurde der Prototyp genutzt, um Informationen zu gewinnen, inwiefern dieser das Nutzererlebnis beim Betrachten von Kurzfilmen beeinflussen konnte. Hierfür wurde entweder der eigene Herzschlag oder der eines Partners durch das künstliche Herz (Tangible-Display), das in der Hand gehalten wurde, ausgegeben. Es zeigte sich, dass eine höhere Immersion auftrat und die Verbundenheit zum Partner verstärkt werden konnte. Diese Ergebnisse könnten daher auch für eine Anwendung bei Meditationsübungen eine Rolle spielen. Insbesondere die höhere soziale Verbundenheit zum Partner stellt zudem eine Verbesserung des sozialen Wohlbefindens dar und könnte mit dem System auch über größere Distanzen ausgelöst werden.

7 Umgebungsbezogenes Wohlbefinden

Umgebungsbezogenes Wohlbefinden beinhaltet alle Auswirkungen der Umgebung auf das Wohlbefinden eines Menschen. In dieser Arbeit werden diesbezügliche Einflüsse innerhalb und außerhalb von Gebäuden behandelt, wobei die Büro- und Wohnumgebung sowie die Stadtlandschaft betrachtet wurden. Die positiven aber auch negativen Einflüsse der Umgebungen können sich auch auf andere Wohlbefindenskategorien auswirken. Beispielsweise beeinflusst ein Spaziergang im Wald das physische (körperliche Bewegung) und mentale (Jennings et al., 2019) Wohlbefinden positiv.

In Abbildung 7.1 ist ein vereinfachter Überblick über die in diesem Kapitel enthaltenen Untersuchungen und deren Bezug zum Wohlbefindensmodell dargestellt.

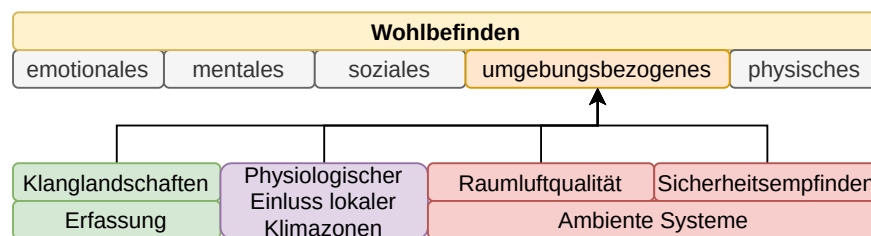


ABBILDUNG 7.1: Überblick über die durchgeführten Untersuchungen in diesem Kapitel.

Erfassung Die Erfassung von Einflüssen der Umgebung kann direkte und indirekte Auswirkungen auf das menschliche Wohlbefinden aufzeigen. Hierbei auftretende wesentliche Probleme und Herausforderungen werden nachfolgend genauer betrachtet.

Problem 1: Erfassung relevanter Umgebungseinflüsse auf den Nutzer

Problem 2: störende, aufwändige Erfassung

Problem 3: Missachtung der Privatsphäre

Am Beispiel einer Audioklassifizierung von Umgebungsgeräuschen wurde folgendes wesentliche Problem beim maschinellen Lernen behandelt:

Problem 4: aufwändige Entwicklung von Netzarchitekturen bei tiefen neuronalen Netzen

Auf den Nutzer wirken in verschiedenen Umgebungen diverse Faktoren, die Einfluss auf das Wohlbefinden besitzen können. Einige relevante Faktoren wurden zur Lösung dieses Problems (**Problem 1**) identifiziert und deren Erfassung wird in diesem Kapitel behandelt (Abschnitt 7.3.1, 7.3.2, 7.3.3, 7.3.4). Zur Analyse von Umgebungsgeräuschen im urbanen Raum wird eine Audioklassifizierung mit tiefen neuronalen Netzen in Abschnitt 7.3.1 beschrieben. Im darauf folgenden Abschnitt 7.3.2 werden Untersuchungen mittels maschinellen Lernen bezüglich der verschiedenen lokalen Klimazonen in der Stadt vorgestellt, für die mobile Geräte (Smartphone, Armband, Umweltsensorik) zur Datenaufzeichnung eingesetzt wurden. Anschließend folgen zwei Arbeiten, die sich mit der Umgebung innerhalb von Gebäuden befassen. In Abschnitt 7.3.3 handelt es sich um die Wohnumgebung, in welcher der Zustand von Fenstern und Türen erfasst wurde, um kontextbezogene, sicherheitsrelevante Erinnerungen auszugeben. Diese Erinnerungen können

positiven Einfluss auf das Sicherheitsgefühl der Hausbewohner besitzen. Die Büroumgebung (Besprechungsraum) wurde in Abschnitt 7.3.4 ausgewählt. Hierbei wurden die Luftqualität gemessen und ausgewertet sowie die Zustände (offen / geschlossen) der Fenster erfasst und ggf. Erinnerungen an das Lüften gegeben. Der Fensterzustand wurde ebenfalls ausgegeben, um beim Verlassen des Raums an das Schließen der Fenster zu erinnern.

Um den Einfluss der Umgebung, in der sich eine Person befindet, mit kontextsensitiven Systemen zu erfassen, ist es nötig, verschiedene Arten von Sensorik zu nutzen (**Problem 2**). In dieser Arbeit wurden daher in den verschiedenen Untersuchungen unaufdringliche tragbare (Wearables) und stationäre (Smarthome) Sensorik eingesetzt. Wearables (Abschnitt 7.3.2) können neben Biosignalen und Bewegungsdaten des Nutzers auch Umgebungsgeräusche ermitteln. Allerdings können sie beispielsweise nicht (zuverlässig) den aktuellen Zustand aller Fenster und Türen in Räumen erfassen, in denen sich eine Person befindet. Auch die Bestimmung verschiedener Gaskonzentrationen ist derzeit mit kompakter, mobiler Sensorik nur begrenzt möglich. Hier kann in die Umgebung integrierte Sensorik (Smarthome (Abschnitt 7.3.3) oder Smartoffice (Abschnitt 7.3.4)) zur Datenerfassung eingesetzt werden.

Der Schutz der Privatsphäre (**Problem 3**) wurde bei sämtlichen Untersuchungen gewährleistet, indem alle Funktionen auf Geräten des Nutzers ausgeführt wurden (Smartphone, Smarthome-Zentrale) oder die eingesetzte Sensorik keine Rückschlüsse auf eine bestimmte Person erlaubte (z. B. Gassensorik, Tür / Fensterkontakte, Bewegungsmelder).

In diesem Kapitel wird anhand der Audioklassifizierung von Umgebungsgeräuschen in Abschnitt 7.3.1 eine Lösung für ein bedeutsames Problem bei der Verwendung von tiefen neuronalen Netzen vorgestellt (**Problem 4**). Die Entwicklung einer gut funktionierenden Netzarchitektur stellt hierbei noch immer eine große Herausforderung dar. In der eigenen Untersuchung wurde hierfür ein genetischer Algorithmus entwickelt und angewandt, der dieses Problem mit akzeptablem Ressourcenaufwand deutlich erleichtern kann.

Informationsausgabe Für die Systeme, die in der Wohn- und Büroumgebung genutzt werden, wurden in dieser Arbeit angepasste unaufdringliche Ausgaben implementiert und evaluiert. Sie folgen der Idee von „Ubiquitous Computing“ bzw. IoT und wurden daher in die Umgebung, als ambiente Anzeigen, unaufdringlich integriert (siehe Abschnitt 7.2). Bei diesen Ausgaben wurden folgende Probleme behandelt:

Problem 1: störende, aufdringliche Informationsausgaben

Problem 2: Unverständlichkeit ambienter Ausgaben

In der Wohn- und Büroumgebung sollte darauf geachtet werden, dass eine unaufdringliche Informationsausgabe stattfindet (**Problem 1**), sodass die Raumnutzer grundsätzlich nicht unnötig gestört bzw. abgelenkt werden. Es muss allerdings ein Kompromiss eingegangen werden, dass die Systemausgaben noch ausreichend Beachtung erhalten. In den beiden Arbeiten in Abschnitt 7.3.3 und 7.3.4 wurde hierbei ambientes Licht für Erinnerungen an offene Fenster / Türen bzw. schlechte Luftqualität und offene Fenster eingesetzt. Dieses Licht war für alle Bewohner bzw. Raumnutzer sichtbar.

Ein wesentliches Problem, das bei ambienten Ausgaben auftreten kann, ist deren Verständlichkeit (**Problem 2**). Eine Lösung hierfür stellt eine personalisierte Kodierung und entsprechende Platzierung dar. Im Fall der ambienten Erinnerung an offene Fenster / Türen in der Wohnung (Abschnitt 7.3.3) wurde bewohnerzentrierte Entwicklung eingesetzt, wodurch die Bewohner die Farbkodierung des ambienten Lichts und dessen Positionierung, bezogen auf deren tägliche Routinen, festlegen konnten. Hierdurch ist ein Verständnis der ambienten Ausgaben gewährleistet. Bei den Erinnerungen an das Lüften in einem Besprechungsraum (Abschnitt 7.3.4) wurden bekannte Farbkodierungen (Ampel) eingesetzt und eine Positionierung neben den Fenstern (verwendet zum Lüften) gewählt. Zusätzlich wurde das Helligkeitsmuster des ambienten Lichts entsprechend dem

Zustand der Fenster angepasst. Hierdurch wurden mehrere Hinweise für den Einsatzzweck der Lichter gegeben.

7.1 Erfassung von Umwelteinflüssen

7.1.1 Umgebungsparameter und deren Auswirkungen auf das Wohlbefinden

Luftqualität

Die Luftqualität kann mittels verschiedener Parameter bestimmt werden. Einige davon besitzen direkten Einfluss auf den menschlichen Körper oder beeinflussen das subjektive Empfinden. In diesem Abschnitt werden die Konzentrationen von Gasen und Feinstaub betrachtet, die mit speziellen Sensoren gemessen oder zumindest abgeschätzt werden können.

Es ist anzumerken, dass für einige Konzentrationen von Luftbestandteilen Empfehlungen der WHO existieren: (*Air quality guidelines: global update 2005: particulate matter, ozone, nitrogen dioxide, and sulfur dioxide*, 2006). Einen aktuellen Überblick über Auswirkungen auf den Menschen zeigt beispielsweise die Arbeit von Schulze et al. (2017).

Kohlendioxid Die Kohlendioxidkonzentrationen (CO_2) wurden im Rahmen dieser Arbeit nur für den Innenbereich betrachtet, da sie dort in höheren Konzentrationen auftreten können. Es werden hier absolute Werte herangezogen, da CO_2 -Konzentrationen auch mit preiswerten Sensoren relativ präzise gemessen werden können, ohne größere Störeinflüsse durch Temperatur / Luftfeuchtigkeit und andere Gase kompensieren zu müssen. Bei absoluten Werten ist zu beachten, dass die mittlere weltweite, atmosphärische CO_2 -Konzentration 2019 bei 409,83 ppm lag und im Mittel, seit 15 Jahren jährlich um 2,24 ppm steigt (siehe Website des Umweltbundesamts¹). Eine Konzentration von ca. 400 ppm ist daher auch in gelüfteten Innenräumen zu erwarten.

Da Menschen beim Atmungsvorgang Sauerstoff aufnehmen und Kohlendioxid ausstoßen, kann die CO_2 -Konzentration genutzt werden, um die allgemeine Luftqualität eines Raumes abzuschätzen. Sie wird daher häufig auch bei sogenannten „Luftqualitätsampeln“ herangezogen. Während der COVID-19-Pandemie, seit Anfang 2020, dienten diese teilweise der Abschätzung, ob Räume gelüftet werden sollten, um eine mögliche Infektionsgefahr durch ausgeatmete infektiöse Aerosole zu verringern².

Die Effekte von Kohlendioxid auf den menschlichen Körper wurden im Literaturüberblick von Azuma et al. (2018) zusammengefasst. Im Folgenden werden einige Arbeiten aus dem Überblick als Quelle herangezogen. Sie dienen dazu, Anhaltspunkte zu erhalten, ab welchen Konzentrationen objektive und subjektive Auswirkungen auf den Menschen messbar sind. Da viele Störfaktoren eine Rolle spielen können, wurden mehrere Studien betrachtet. Diese nutzten meist eine kontrollierte Umgebung, in der die Kohlendioxidkonzentration gesteuert werden konnte, und verwendeten standardisierte Tests zur Bestimmung der kognitiven Leistungsfähigkeit bzw. Fragebogen zur Ermittlung der subjektiv empfundenen Luftqualität.

In einer kontrollierten Studie von Satish et al. (2012) konnte eine Reduktion der kognitiven Leistungsfähigkeit ab einer CO_2 -Konzentration von 1000 ppm nach 2,5 Stunden beobachtet werden. Allen et al. (2016) konnten in ihrer Laborstudie eine Verringerung der Leistungsfähigkeit bei 950 ppm feststellen. In einer anderen Laborstudie von Maula et al. (2017) wurde bei 2260 ppm nach 4 Stunden eine reduzierte kognitive Leistung gemessen. Die von den Probanden wahrgenommene Arbeitslast und Müdigkeit war bei dieser Konzentration erhöht. In der kontrollierten Studie von X. Zhang et al. (2017) konnte allerdings keine signifikante Reduktion der kognitiven Leistungsfähigkeit bis zu einer Konzentration von 3000 ppm über 4,25 Stunden festgestellt werden.

¹<https://www.umweltbundesamt.de/daten/klima/atmosphaerische-treibhausgas-konzentrationen> (10.09.2020)

²<https://www.br.de/nachrichten/wissen/frische-luft-im-klassenzimmer-was-sind-co2-ampeln,S9ysgT9> (31.03.2021)

Sobald Bio-Effluente (menschliche Atmung, Körpergerüche) ebenfalls in der Luft vorhanden waren, war die kognitive Leistung jedoch verringert. Die Studienteilnehmer bewerteten die Luftqualität schlechter und berichteten von verstärkter Intensität von Kopfschmerzen, Müdigkeit, Schläfrigkeit und erhöhter Schwierigkeit, klar zu denken. Die Autoren vermuteten, dass sie keine Abhängigkeit zwischen kognitiver Leistungsfähigkeit und CO₂-Konzentration feststellen konnten, da sie einfachere kognitive Tests verwendeten. Allen et al. (2016) und Satish et al. (2012) nutzten beispielsweise den komplexeren „strategic management simulation“ Test.

In einer aktuelleren Studie von Allen et al. (2018) wurde die Leistung von Piloten kommerzieller Fluggesellschaften untersucht, indem drei Stunden lange Flüge mit veränderten CO₂-Konzentrationen durchgeführt wurden. Insbesondere bei einer Konzentration von 2500 ppm war die Leistung der Piloten signifikant geringer als bei 700 ppm. Bei 1500 ppm konnte nur eine geringe Leistungsreduktion festgestellt werden.

Einstufungen verschiedener Kohlendioxidkonzentrationen können auch in offiziellen Dokumenten gefunden werden. Beispielsweise in der Bekanntmachung „Gesundheitliche Bewertung von Kohlendioxid in der Innenraumluft“ (2008) des Umweltbundesamts. Hier wurden mehrere Studien vor 2008 berücksichtigt. So werden Konzentrationen unter 1000 ppm als unauffällig, zwischen 1000 ppm und 2000 ppm als auffällig und über 2000 ppm als inakzeptabel eingestuft.

Insgesamt können die Bewertungsstufen der Luftqualität in Innenräumen des Umweltbundesamts auch heute noch genutzt werden, da neuere Studien keine relevanten Unterschiede zu den angegebenen Schwellwerten zeigen. Sowohl objektive als auch subjektive negative Einflüsse auf den Menschen sind gemäß den Studien bei höheren CO₂-Konzentrationen zu erwarten.

VOCs und andere Gase Flüchtige organische Verbindungen (VOCs) bestehen aus verschiedenen Gasen und werden in vielen Fällen zur Bestimmung der Luftqualität herangezogen. In der Raumluft können diese laut Rivas et al. (2019) beispielsweise von folgenden Quellen stammen: Reinigungsmittel, Zigarettenrauch, Baumaterialien (z. B. Laminat, Farben, Klebstoffe) und Computern. Es sind allerdings auch Gerüche erfassbar, die beim Kochen entstehen oder von Menschen stammen.

Einige der Gase sind für den Menschen in höheren Konzentrationen schädlich oder beeinträchtigen ihn temporär. Es existieren verschiedene Sensoren, die auf VOCs reagieren können. Diese erfassen meist auch Kohlendioxid. Kostengünstige Sensoren, liefern jedoch Werte, die sich auf eine Referenz-Konzentration (z. B. „frische Luft“ außerhalb von Gebäuden) beziehen. Die zurückgegebenen Absolutwerte sind daher nur begrenzt aussagekräftig. Zudem kann mit dieser Sensorik nicht bestimmt werden, um welche Gase es sich handelt. In der Arbeit von Allen et al. (2016) bezüglich der Auswirkungen von CO₂- und VOCs-Konzentration auf die kognitive Leistungsfähigkeit von Menschen wurden die VOCs-Konzentrationen nachträglich in einem Labor mittels gesammelter Gasproben untersucht und die Einzelgase aufgeschlüsselt. Um zu erkennen, ob gelüftet werden sollte, ist eine Messung mit einfachen VOCs-Sensoren in der Regel ausreichend. Im Vergleich zu CO₂-Sensoren können diese zusätzlich unangenehme oder gar schädliche Gase erfassen.

Insbesondere in Städten spielen bei der Beurteilung der Luftqualität noch weitere Gase eine Rolle, die auch mit einfachen Sensoren gemessen werden können. Hierzu zählen beispielsweise Stickstoffdioxid (NO₂) und Kohlenstoffmonoxid (CO), die u. a. bei Verbrennungsmotoren von Fahrzeugen freigesetzt werden. Diese können durch offene Fenster oder Belüftungen auch in Innenräume gelangen. Kurzfristige Exposition gegenüber NO₂ führte in einer Studie von Brand et al. (2016) nicht zu gesundheitlichen Auswirkungen bei jungen Erwachsenen, allerdings wurden in der Arbeit von Gaffin et al. (2018) negative Effekte bei Schulkindern mit Asthma berichtet.

Beim nicht wahrnehmbaren CO ist bekannt, dass hohe Konzentrationen massive gesundheitsschädliche bzw. tödliche Auswirkungen besitzen. Laut dem Artikel von Bosch (2016) wird die Fähigkeit des Blutes Sauerstoff zu transportieren reduziert. Zu den unspezifischen Symptomen

einer leichten Kohlenmonoxidvergiftung zählen gemäß dem Autor beispielsweise Schwindel, Kopfschmerzen, Desorientierung und Übelkeit. Im Artikel werden als Kohlenmonoxidquelle unvollständigen Verbrennungsprozesse genannt, die z. B. bei Verbrennungsmotoren, Gasthermen, offenen Feuerstellen und (Wasserpfeife)-Rauchen auftreten. Insbesondere in geschlossenen oder schlecht gelüfteten Räumen kommt es häufig zu schweren Unfällen durch hohe CO-Konzentrationen. Zum Schutz vor Vergiftungen werden daher beispielsweise in Tiefgaragen Warngeräte installiert. Auch für den privaten Haushalt sind preisgünstige Sensoren verfügbar.

Feinstaub Feinstaub kann relativ präzise mit kostengünstiger Sensorik gemessen werden, wobei die absoluten Partikelkonzentrationen in $\mu\text{g}/\text{m}^3$ für $\text{PM}_{2,5}$ und PM_{10} für die Bewertung der Luftqualität bestimmt werden. PM steht hierbei für „particulate matter“. Bei $\text{PM}_{2,5}$ ist der Durchmesser der Partikel kleiner als $2,5\ \mu\text{m}$, bei PM_{10} liegt der Durchmesser unter $10\ \mu\text{m}$. Laut dem Artikel von Schulze et al. (2017) können größere Partikel vom Körper gefiltert werden, sind sie allerdings kleiner als $10\ \mu\text{m}$ erreichen sie die Lungenbläschen. Vom Menschen produzierter Feinstaub stammt, gemäß den Autoren, beispielsweise aus Verbrennungsmotoren oder Zigarettenrauch. Einige der im Artikel betrachteten Studien zeigten, dass die Krebsrate mit einer Erhöhung der Feinstaubkonzentration proportional ansteigt, zudem kann Feinstaub diverse schwerwiegende negative Effekte auf schwangere Frauen und deren ungeborene Kinder besitzen. Andere Quellen der Arbeit deuten darauf hin, dass Asthma, Atemwegserkrankungen, Lungenkrebs und Herz-Kreislauf-Erkrankungen verursacht werden können.

Temperatur und Luftfeuchtigkeit

Die Umgebungstemperatur und Luftfeuchtigkeit spielen eine bedeutsame Rolle für das Wohlbefinden von Menschen. Häufig werden hierfür auch Luftströmungen oder Strahlungswärme berücksichtigt, die allerdings nur mit spezieller, meist teurer Sensorik messbar sind.

Die in der Thermodynamik definierten Begriffe Wärme, Temperatur und thermische Energie spielen hier eine Rolle. Im Buch von Hering et al. (2016, S. 170 f.) werden die Begriffe folgendermaßen definiert: „Wärme ist Energie, die aufgrund eines Temperaturunterschieds zwischen zwei Systemen übertragen wird. Diese Energieübertragung hat eine eindeutige Richtung. Die Wärme fließt stets in Richtung der niedrigeren Temperatur. Der Wärmeübergang ist also ein irreversibler Prozess.“. Der Unterschied zwischen der Umgebungstemperatur und der Körpertemperatur bestimmt daher, wie viel Wärme vom menschlichen Körper an die Umgebung abgegeben (Körper ist wärmer) oder von der Umgebung aufgenommen (Körper ist kühler) wird. Bei deutlich wärmerer Umgebungstemperatur entsteht laut dem Artikel von McGregor und Vanos (2018) Wärmestress und die Haut beginnt damit, Schweiß zu produzieren, wobei sich auch die Herzrate erhöht. Gemäß dem Artikel spielt die Luftfeuchtigkeit insofern eine Rolle, als eine geringe relative Luftfeuchtigkeit es dem Körper erlaubt, sich besser durch Schwitzen zu kühlen. Bei hoher Luftfeuchtigkeit ist dies allerdings ineffizient. Gemäß den Autoren des Artikels kann eine Belüftung bis zur Hauttemperatur von ca. $36\ ^\circ\text{C}$ kühlen, solange die Lufttemperatur unterhalb dieser Temperatur liegt. Zur Vorhersage der Wärmebelastung auf den Menschen existieren verschiedene Indizes, wie beispielsweise der von den Autoren behandelte „Heat index“, der in den USA für Warnungen vor Hitzewellen verwendet wird. Dieser Index kann mit relativ geringem Aufwand mittels Temperatur und relativer Luftfeuchtigkeit bestimmt werden.

Negative Auswirkungen auf den Menschen treten sowohl in zu warmer als auch in zu kalter Umgebung auf. Hierzu wurden im Literaturüberblick von Martin et al. (2019) die kognitiven Auswirkungen von Hitze und Kälte ausführlich behandelt. Die Studien nutzten unterschiedliche kognitive Aufgaben und es zeigte sich, dass bei erhöhter Wärmeeinwirkung die Fehlerquoten und Reaktionszeiten der Probanden häufig anstiegen. Zudem verringerte sich deren Merkfähigkeit und Genauigkeit bei der Aufgabenbearbeitung. Dies war bedingt auch bei kalten Umgebungen zu beobachten. Generell werden in Innenräumen gemäß (Neufert & Neufert, 2012, S. 37) folgende

Lufttemperaturen als komfortabel empfunden: im Sommer zwischen 20 und 24 °C, im Winter ca. 21 °C (± 1). Die Temperatur der umgebenden Oberflächen sollte hierbei nicht mehr als 2 bis 3 °C von der Lufttemperatur abweichen. Laut dieser Quelle sollte die Luftfeuchtigkeit bei 40 bis 50 % liegen. Bei der wahrgenommenen Wärme im Außenbereich spielt die Vegetation eine erhebliche Rolle. Beispielsweise konnte im Artikel von Louafi et al. (2017) eine Beziehung zwischen der Vegetation und dem thermischen Wohlbefinden in heißem Klima festgestellt werden. Dies wird unter anderem dadurch verbessert, dass Bäume Schatten spenden und die Luftfeuchtigkeit über ihre Blätter erhöhen.

Im Innenbereich können erhöhte Luftfeuchtigkeit und ungünstige Temperaturen zusätzliche negative Auswirkungen auf die Gesundheit und das Wohlbefinden verursachen, da hierdurch das Wachstum von Pilzen und Schimmel begünstigt wird. Diese können u. a. Allergene, Gifte und VOCs produzieren, die beispielsweise zu toxischen Reaktionen, Irritationen, Infektionen und Allergien führen können (K.-H. Kim et al., 2018). Entsprechende wachstumsfördernde Umgebungen, wie sie nach dem Duschen / Baden im Bad auftreten, können jedoch durch Überwachung von Temperatur und Luftfeuchtigkeit häufig reduziert oder vollständig unterbunden werden.

Neben den bereits erwähnten gesundheitlichen Risiken durch Schimmelpilze kann das Wohnen in einer kalten, feuchten Wohnung laut Liddell und Guiney (2015) zusätzlich zu einem reduzierten mentalen Wohlbefinden führen, da verschiedene Stressoren auftreten. Neben der Angst vor gesundheitlichen Schäden zählen gemäß dem Artikel hierzu auch die Befürchtungen hinsichtlich einer Beschädigung der Bausubstanz und von Einrichtungsgegenständen. Auch soziale Isolation kann sich durch eine unansehnliche Wohnung ergeben.

Landschaft

Bereits im vorherigen Abschnitt wurden die Auswirkungen von Temperatur und Luftfeuchtigkeit auf die Gesundheit und das menschliche Wohlbefinden beleuchtet, die unter anderem auch von der Landschaft abhängen. Insbesondere in Städten spielen hierbei Grünflächen wie z. B. Parks eine wichtige Rolle. Vor- und Nachteile solcher urbaner Grünflächen werden nachfolgend, basierend auf dem Literaturüberblick von Jennings et al. (2019), aufgezeigt.

Grünflächen können einen positiven Effekt auf die mentale Gesundheit, kognitiven Funktionen und den Stresslevel besitzen. Die Reduktion des Stresslevels besitzt weitere vorteilhafte Auswirkungen. Durch dauerhaften hohen Stresslevel wird beispielsweise das Immunsystem geschwächt, ein erhöhtes Risiko für Diabetes entsteht, kardiovaskuläre Krankheiten können ausgelöst werden und die Sterblichkeit erhöht sich. Ein positiver Effekt auf die mentale Gesundheit und kognitiven Funktionen konnte mittels funktioneller Magnetresonanztomografie (fMRI) im Labor nachvollzogen werden. In Studien ist beispielsweise eine Reduktion depressiver Symptome durch Spaziergänge festgestellt worden. Indirekte gesundheitliche Vorteile werden dadurch ermöglicht, dass Grünflächen dazu einladen, körperliche Aktivitäten durchzuführen. Hierdurch kann zudem die soziale Interaktion gefördert werden, wenn beispielsweise zusammen gewandert oder eine Gruppensportart betrieben wird. Bereits ein Aufenthalt im Wald kann das Immunsystem verbessern, da hier bestimmte gesundheitsfördernde Pflanzen und Bakterien leben. In Studien konnte zudem eine erhöhte Aktivität von Antikrebszellen beobachtet werden. Es gibt Hinweise darauf, dass bereits die Sicht auf Grünflächen positive Effekte besitzt. In Studien konnte festgestellt werden, dass sich beispielsweise Krankenhauspatienten um 15 % schneller erholten, wenn sie in Zimmern lagen, die Ausblick auf Grünflächen besaßen, im Vergleich zu Räumen, wo nur eine Ziegelmauer sichtbar war.

Allerdings bestehen laut Jennings et al. (2019) noch einige Unklarheiten über die genauen positiven gesundheitlichen Auswirkungen von Grünflächen auf den Menschen. Beispielsweise ist bislang nicht hinreichend geklärt, welche Art von Natur am effizientesten wirkt und welche Dauer und Häufigkeit von Aufenthalten in der Natur notwendig ist. Allerdings besitzen Grünflächen nicht nur Vorteile. Ein Problem besteht darin, dass dort Pflanzen wachsen können, deren

Pollen bei einigen Menschen allergische Reaktionen auslösen. Zur Vermeidung dieser möglichen gesundheitlichen Beeinträchtigung sollten diesbezüglich unproblematische Pflanzensorten zur Begrünung ausgewählt werden. Weitere Gesundheitsrisiken können entstehen, falls Grünflächen schlecht gestaltet und gepflegt sind. Es können beispielsweise eine große Anzahl an Stechmücken auftreten, wenn natürliche Fressfeinde, wie z. B. verschiedene Vogelarten, keinen Unterschlupf finden. Zudem können sich Grünflächenbesucher verletzen, falls Bäume nicht regelmäßig von toten Ästen befreit werden und diese auf Wege fallen.

Es ist ersichtlich, dass die Vorteile die Nachteile deutlich überwiegen. Das Messen dieses Potenzials zur Verbesserung des menschlichen Wohlbefindens gestaltet sich allerdings schwierig, da viele Faktoren eine Rolle spielen, die teilweise wie z. B. die Ästhetik nicht direkt messbar sind und möglicherweise erst in Kombination ihre volle Wirkung entfalten. Relativ unproblematisch messbare Parameter wurden zum Teil in der vorliegenden Arbeit bereits behandelt. Hierzu zählen die Luftqualität, Temperatur und Luftfeuchtigkeit. Visuell kann Pflanzenbewuchs beispielsweise mittels Fotos erfasst und analysiert werden. Hierbei ist allerdings nicht bekannt, welche Arten und Sorten von Pflanzen besonders vorteilhaft auf den Menschen wirken, da dies auch von der subjektiven Wahrnehmung abhängt.

Klanglandschaft

Zur Landschaft gehört zusätzlich die auditive Dimension, die in diesem Abschnitt genauer behandelt wird. Die Klanglandschaft (engl. Soundscape) ist Teil der Umgebungswahrnehmung und liefert wesentliche Informationen darüber, in welcher Umgebung (z. B. neben einer Straße) sich eine Person befindet oder welche Aktivitäten sie ausführt. Dies kann indirekte Rückschlüsse auf das potenzielle Wohlbefinden einer Person erlauben.

Der Standard DIN ISO 12913-1 ist eine neuere Bemühung, den Begriff „Soundscape“ international festzulegen, wobei er folgendermaßen definiert wird: „akustische Umgebung, die durch eine Person oder durch eine Gruppe von Menschen im Kontext wahrgenommen, erfahren und / oder begriffen wird“ (Fiebig & Schulte-Fortkamp, 2019). Das Forschungsgebiet „Soundscape“ überschneidet sich laut Aletta und Xiao (2018) mit verschiedenen anderen Disziplinen, wie z. B. Soziologie, Umweltpsychologie, Musik, akustischen Ökologie, Stadtplanung, Lärmschutzbau und Architektur. Pioniere im Bereich Soundscapes sind Southworth und Schafer, die insbesondere die Frage stellten, wie moderne Städte sich anhören und nicht nur wie sie aussehen sollten (Aletta & Xiao, 2018). Laut Aletta und Xiao (2018) können Soundscapes in nahezu allen Kontexten untersucht werden, allerdings ist insbesondere der urbane Kontext in der Forschung von Bedeutung, da hier besondere Herausforderungen durch die zunehmende Verstädterung und Globalisierung auftreten.

Eine Klanglandschaft kann laut Schafer (1993) durch folgende Eigenschaften charakterisiert werden: „keynote sounds“, „signals“ und „soundmarks“. Die Begriffe werden nach Schafer (1993), im Folgenden definiert. *Keynote sounds* bestimmen den Hintergrund und werden nicht immer bewusst wahrgenommen. In einer Landschaft sind dies beispielsweise Geräusche von Wasser, Wind, Vögeln und Insekten. *Signals* sind im Vordergrund und werden bewusst wahrgenommen. Einige davon dienen zur Warnung, wie beispielsweise Sirenen oder Hupen. Der Begriff *soundmark* ist von „landmark“ (deutsch Landmarke) abgeleitet. Dieses Geräusch ist für bestimmte Bereiche einzigartig oder es wird von Personen in diesem Bereich besonders beachtet. Dies kann beispielsweise eine Kirchturmuhren mit Glockenspiel oder aber auch ein Nebelhorn sein. Klanglandschaften können laut Schafer (1993) zusätzlich bezüglich ihres Dynamikbereichs in „hi-fi“ und „low-fi“ unterteilt werden. In einer *hi-fi* Klanglandschaft können einzelne Geräusche klar bestimmt werden, da nur ein geringer Hintergrundgeräuschpegel vorhanden ist. Geräusche überlappen zudem seltener. Dies ist laut dem Autor beispielsweise im ländlichen Bereich eher der Fall, als in der Stadt. In der Stadt herrscht meist eine *lo-fi* Klanglandschaft, mit einer hohen Dichte an überlagerten Geräuschen. Die hiermit eingeführten, generellen Begriffe zur Beschreibung von Klanglandschaften helfen

bei der Einordnung von Geräuschen in der Umgebung und spielen auch bei der automatischen Klassifizierung eine Rolle.

Literatur zu positiven Gesundheitsauswirkungen von Klanglandschaften wurden im Überblick von Aletta et al. (2018) zusammengestellt. In einigen Studien konnte beispielsweise beobachtet werden, dass angenehme Klanglandschaften wie Naturgeräusche es erlauben, Stress schneller abzubauen. Stille Soundscapes können zudem das physiologische und psychologische Wohlbefinden verbessern. Dies wurde ebenfalls von R. T. Buxton et al. (2021) in einer Literaturanalyse bei 18 Quellen festgestellt. Gemäß dieser können natürliche Geräusche die Gesundheit verbessern, positiven Affekt erhöhen sowie Stress und Ärger verringern. Die Lautstärke von Geräuschen spielt generell eine erhebliche Rolle und kann zusätzlich gemessen werden. Hierbei ist diese von ihrer Intensität her messbar und es sind Grenzwerte bekannt, bei denen gesundheitliche Schäden am Ohr auftreten. Bei einem Schalldruckpegel über 130 dB(A) kommt es zu einem Hörschaden (Schalltrauma), dauerhafte Pegel von über 90 dB(A) können Schäden am Gehör verursachen (Schmidt & Thews, 2013). Eine subjektive Lärmbelästigung kann laut Schmidt und Thews (2013) einen Anstieg des Blutdrucks sowie Schlaflosigkeit verursachen. Bei solchen Geräuschen mit geringeren Schalldruckpegeln sind Auswirkungen allerdings individuell.

7.1.2 Eigene Arbeiten

Eine Erfassung von Bestandteilen einer Klanglandschaft wurde in Abschnitt 7.3.1 untersucht. Diese wurde im Rahmen der DCASE-Challenge 2018 entwickelt, wodurch die Ergebnisse mit anderen Teilnehmern vergleichbar sind. Die Besonderheit der eigenen Arbeit lag darin, dass hierbei ein genetischer Algorithmus die verwendeten Netzarchitekturen von tiefen neuronalen Netzen erstellt und ausgewählt hat, die damit deutlich besser abschnitten, als die per Hand entwickelten Netze der Baseline der Veranstalter. Das Verfahren konnte dieses Ergebnis zudem mit mehreren normalen Arbeitsrechnern (CPU: Intel i7-6700, RAM: 16 GiB) mit mittelmäßig leistungsfähigen GPUs (NVIDIA GTX 1060, RAM: 6 GiB) nach einer Laufzeit von ca. 5 Tagen erreichen. Bessere Ergebnisse wären daher mit weiteren Optimierungen am Algorithmus und der restlichen verwendeten Software möglich gewesen. Das Verfahren ist generisch und somit auch für andere Klassifizierungsprobleme anwendbar.

In Abschnitt 7.3.2 wurde untersucht, ob mithilfe tragbarer Sensorik Auswirkungen verschiedener lokaler Klimazonen in urbaner Umgebung erfasst und diese auch mittels Daten eines Fitnessarmbands (Herzrate, IBI, Hautleitwert) erkannt werden können. Hierfür wurden mehrere Aufzeichnungen an verschiedenen Tagen über eine festgelegte Strecke mit unterschiedlichen Personen durchgeführt. Die Strecken enthielten drei verschiedene lokale Klimazonen, die mittels der Biosignale des Fitness-Trackers klassifiziert werden konnten. Zusätzlich konnte gezeigt werden, dass die emotionale Valenz (negativ, positiv) mit den Biodaten und auch (weniger genau) mit Audiodaten des Smartphones vorhergesagt werden kann. Die emotionale Valenz spielt beim emotionalen Wohlbefinden eine bedeutsame Rolle (siehe Kapitel 6). Für zukünftige Anwendungen wurden die Verfahren so ausgewählt, dass sie lokal, mittels mobileSSI (siehe Abschnitt 4.1.2) für eine Klassifizierung in Echtzeit geeignet sind.

Bei den Arbeiten in Abschnitt 7.3.3 und 7.3.4 handelt es sich um vollständig entwickelte Systeme in der Wohn- bzw. Büroumgebung (Besprechungsraum). Hierbei wurden Sensordaten ambienter Sensorik, die von einem Smarthome-System erfasst werden können (Nutzerpräsenz, Zustand von Türen / Fenstern und Luftqualität) aufbereitet und per ambientem Licht ausgegeben. Es wurde hierfür entsprechende Smarthome-Software eingesetzt bzw. eigene Software (Eventinterpretor (siehe Abschnitt 4.2.2)) zur Datenverarbeitung entwickelt. Ein besonderes Alleinstellungsmerkmal im Vergleich zu anderen Arbeiten stellt bei diesen Systemen jedoch die Systemausgabe per ambientem Licht dar und wird im Rahmen der Informationsausgabe detaillierter behandelt.

7.2 Ambiente Systeme

IoT, ambiente Sensorik und Ausgaben Ambiente Sensorik und ambiente Anzeigen (engl. ambient displays) können als Teil stiller Technologie (engl. calm technology) angesehen werden. Dieser Begriff wurde von Mark Weiser (Weiser & Brown, 1995) geprägt und steht für Technologie, die normalerweise nur im Hintergrund wahrgenommen wird, sich allerdings auch nahtlos in den Vordergrund der Aufmerksamkeit begeben kann. Sie zählt zum „Ubiquitous Computing“, das ebenfalls von Mark Weiser eingeführt und für das 21. Jahrhundert vorhergesagt wurde (Weiser, 1991). Das Internet der Dinge (IoT) basiert auf dieser Idee und setzt diese mit aktuellen Technologien um. Vernetzte Systeme arbeiten hierbei primär im Hintergrund und erfassen mit unaufdringlicher, ambienter Sensorik benötigte Daten und können Systemausgaben wiederum ohne zu stören ausgeben. Sensoren aus dem Smarthome-Bereich, wie beispielsweise Bewegungsmelder, Fensterkontakte oder Gassensoren sind Teil des Internets der Dinge und werden unauffällig in den Wohnraum integriert. Siehe hierzu auch Kapitel 3. Unaufdringliche Ausgaben können mittels ambienter Anzeigen realisiert werden. Gemäß der Definition von ambienten Anzeigen, die unter anderem in der Arbeit von Mankoff et al. (2003) genutzt wurde, sollten diese peripher kommunizieren und hierbei minimale Aufmerksamkeit sowie kognitive Last benötigen. Ambiente Ausgaben stellen daher eine ideale Kombination zu den sonstigen im Hintergrund arbeitenden IoT-Geräten dar.

Implizite und explizite Systeminteraktion Insbesondere bei der Interaktion mit IoT-Geräten tritt ein Wechsel zwischen Interaktion im Hintergrund und Vordergrund auf. Bei der Entwicklung von Systeminteraktionen kann das Vordergrund- / Hintergrund- Interaktionsmodell von B. Buxton (1995) eingesetzt werden. In diesem Modell ist neben der Mensch-Computer-Interaktion auch die Mensch-Mensch-Interaktion enthalten. Letztere wird im Folgenden allerdings nicht weiter behandelt, da sich diese Arbeit primär mit Mensch-Computer-Interaktion befasst. Aktionen im Vorder- und Hintergrund beziehen sich hierbei auf die Aufmerksamkeit des Menschen. Das manuelle, explizite Ein- und Ausschalten von Licht mit einem Schalter ist beispielsweise eine bewusste Aktion, die im Vordergrund stattfindet. Wenn das Licht mit einem Bewegungsmelder selbstständig, implizit gesteuert wird, handelt es sich um eine Interaktion, die im Hintergrund abläuft. Das System versucht hierbei, die Intention des Nutzers vorherzusehen und die entsprechende Aktion selbstständig durchzuführen. Falls ein automatisches System auch eine manuelle Licht-Steuerung bzw. Übersteuerung per Schalter zulässt, kann die Interaktion zwischen Vordergrund und Hintergrund wechseln. Dies kann notwendig sein, wenn das System beispielsweise wegen fehlender Sensorik oder Logik nicht erkennt, dass das Licht ausgeschaltet werden kann, da von außen ausreichend Sonnenlicht in den Raum fällt. Im Interaktionsmodell von Ju und Leifer (2008) („Implicit Interaction Framework“) wurde die Dimension „Initiative“ ergänzt, mit der bestimmt werden kann, ob sich ein System eher reaktiv oder proaktiv verhält. Ein reaktives System reagiert auf Nutzeraktionen, ein proaktives führt Aktionen automatisch aus. Verschiedene Interaktionen mit einem System können somit mittels der benötigten Nutzeraufmerksamkeit (Vorder- bzw. Hintergrund) sowie der Systeminitiative (reaktiv bzw. proaktiv) beschrieben werden.

Modalitäten ambienter Ausgaben In ihrer Arbeit betonen Mankoff et al. (2003), dass ambiente Anzeigen nicht auf die visuelle Wahrnehmung beschränkt sind. Hieraus folgt, dass ambiente Ausgaben grundsätzlich für alle menschlichen Sinne entwickelt werden können. Basierend auf den klassischen fünf Sinnen existieren somit auditive, gustatorische, haptische, olfaktorische und visuelle Anzeigen. Die Nutzung der verschiedenen Modalitäten als Systemausgaben besitzt verschiedene Vor- und Nachteile und sollte, je nach Anwendungsfall, ausgewählt bzw. kombiniert werden. **Gustatorische** Ausgaben sind beispielsweise mit aktueller Technik nicht ambient umsetzbar, da hierfür die Zunge stimuliert werden muss, wofür Elektronik oder chemische Flüssigkeiten dieser zugeführt werden müssen (siehe Überblick über gustatorische Interfaces: Vi et al. (2017)). Diese

technischen Einschränkungen erlauben zudem nur eine individuelle Nutzung. **Olfaktorische** Ausgaben (siehe Überblick von Efe (2017)) gehören ebenfalls zu den unüblichen Ausgabeformen. Ein wesentlicher Nachteil dieser Modalität besteht darin, dass meist nur wenige ausgewählte Gerüche erzeugt werden und sich diese durch Luftströmungen in Räumen ungünstig verteilen können. Zudem kann hierdurch eine sehr langsame Kommunikation auftreten, die durch zuvor eingesetzte oder fremde Gerüche gestört werden kann. Bereits eine normale Erkältung oder Heuschnupfen können dafür sorgen, dass Gerüche nicht mehr wahrgenommen werden. **Haptische** Ausgaben werden häufig bei mobilen Geräten, direkt am Körper, für persönliche Meldungen eingesetzt. Gezielte Luftbewegungen mit Lüftern oder Ultraschall (Ablart et al., 2019), vibrierende Böden oder thermische Haptik (Chen et al., 2017) können jedoch auch in einem Raum ein spürbares Feedback erzeugen. Ein Nachteil ist hierbei der im Vergleich zu auditiven oder visuellen Displays hohe technische Aufwand. Eine **auditive** ambiente Anzeige wird beispielsweise in der Publikation von Yu et al. (2017) präsentiert. Bei dieser wurden anpassbare, natürliche Klanglandschaften zur Informationsausgabe eingesetzt. Die auditive Modalität besitzt ähnliche Probleme wie die visuelle. Hierzu gehört beispielsweise, dass taube bzw. blinde Personen diese nicht nutzen können, wenn ausschließlich der entsprechende Sinn eingesetzt wird. Wie auch bei den anderen Ausgabeformen können Überlagerungen stattfinden. Insbesondere eine Überlagerung der menschlichen Sprache ist allerdings bei der auditiven Modalität zu vermeiden, da dies die wichtigste Kommunikationsform zwischen Menschen darstellt. **Visuelle** ambiente Anzeigen werden am häufigsten eingesetzt, da dies beim Menschen den ausgeprägtesten Sinn darstellt. Diese können mittels Bildern auf Displays (siehe hierzu digitale Bilderrahmen in Abschnitt 9.1.2), Projektionen bzw. Schattenspiele (z. B. (Jensen et al., 2019)), physischen Objekten (z. B. (Sauvé et al., 2017)) und Licht (siehe eigene Arbeiten) umgesetzt werden. Es existieren auch Kleidungs-Prototypen mit integrierten ambienten Anzeigen (z. B. (Harjuniemi et al., 2020)).

Eigene Arbeiten Bei den eigenen Systemen aus Abschnitt 7.3.3 und 7.3.4 wurde Smarthome-Technologie sowie als Ausgabemedium ambientes Licht eingesetzt. Die CARE-Prototypen in Kapitel 9 besaßen einen ambienten Modus, bei dem das eingesetzte Display sich wie ein digitaler Bilderrahmen verhielt und das Interesse der Nutzer wecken sollte. Spätere Prototypen wurden um ein ambientes Licht erweitert. Ambientes Licht wurde für alle Systeme als Modalität ausgewählt, da es vergleichsweise unproblematisch und kostengünstig in vorhandene Räume integriert und die Lichtquelle gezielt neben bestimmten Objekten platziert werden kann. Zudem können verhältnismäßig viele verschiedene Informationen mittels Farbe und Helligkeitsmustern kodiert werden. Des Weiteren ist es insbesondere für Senioren einfacher, die Farben und Helligkeitsmuster wahrzunehmen, da beispielsweise der Hörsinn im Alter häufiger und weitreichender von Einschränkungen betroffen ist.

Die in diesem Kapitel in Abschnitt 7.3.3 und 7.3.4 vorgestellten Systeme für die Wohn- bzw. Büroumgebung setzten ambientes Licht mit Farb- und Helligkeitskodierung ein. Das Problem des fehlenden Bezugs bei ambientem Licht wurde mittels der Platzierung des Lichtes gelöst. In der Wohnung wurden sie so platziert, dass sie den Alltagsroutinen entsprachen, sodass beispielsweise beim Verlassen der Wohnung ein ambientes Licht in unmittelbarer Nähe zur Haustür anzeigte, ob Fenster / Türen noch offen waren oder alles „sicher“ war. Bei der Erinnerung an schlechte Luftqualität in einem Besprechungsraum wurden die verwendeten LED-Streifen neben den Fenstern platziert, die zum Lüften geöffnet werden konnten. Die Nutzung von ambientem Licht mit gezieltem Einsatz der Lokalität der Leuchtmittel zur Anzeige spezieller, relevanter und ansonsten nicht direkt sichtbarer Daten (Zustand der Fenster / Türen und Luftqualität) der Systeme war in dieser Kombination neuartig.

Eine Einordnung der Nutzerinteraktion mit den Systemen ist mittels des „Implicit Interaction Framework“ von Ju und Leifer (2008) möglich. Die zuvor genannten Systeme in dieser Arbeit beinhalten implizite Nutzerinteraktion. Die Systeme in den Abschnitten 7.3.3 und 7.3.4 agieren hierbei

ausschließlich im Hintergrund und arbeiten proaktiv, da sie mit in der Wohn- bzw. Büroumgebung integrierter Sensorik Daten erfassen, auswerten und mit ambientem Licht, unaufdringlich ausgeben. Auf Nutzeraktionen wurde hierbei ausschließlich implizit reagiert, falls der Nutzer beispielsweise den Zustand der Fenster änderte.

Beim CARE-System handelte es sich um ein System, das kontextsensitiv Empfehlungen zur Verbesserung des Wohlbefindens von Senioren auswählte. Hierbei wurde Smarthome-Sensorik eingesetzt und ein Tablet in die Wohnumgebung integriert. Das Tablet wurde so verwendet, dass es wie ein digitaler Bilderrahmen agierte, wenn keine Person direkt vor dem System stand. Wenn sich im Raum eine Person befand, wurde ein ambientes Licht mit einer Farbkodierung aktiviert. In diesem Zustand agierte das System proaktiv im Hintergrund. Sobald sich die Person jedoch näherte, wurden Empfehlungen angezeigt und der Nutzer konnte explizit über die GUI auf dem Touchscreen oder Taster interagieren. In diesem Modus war das System somit proaktiv im Vordergrund bzw. reaktiv, wenn der Nutzer Eingaben tätigte. Sobald die Person sich vom System entfernte wurde wiederum der Bilderrahmen-Modus aktiviert.

7.3 Eigene Untersuchungen

7.3.1 Klassifizierung von Klanglandschaften mit generierten DNNs

Einleitung

Umgebungsgeräusche können einen großen Einfluss auf das Wohlbefinden des Menschen besitzen. Diese Geräusche erlauben direkte Rückschlüsse darauf, ob die Umgebung als angenehm oder störend empfunden wird. Falls beispielsweise ein Presslufthammer in der Nähe verwendet wird, ist anzunehmen, dass einerseits die Lautstärke und andererseits die Art des Geräusches als sehr unangenehm empfunden werden. Zudem kann anhand von Geräuschen der Aufenthaltsort von Personen grob bestimmt werden. Dies ermöglicht einen Rückschluss auf deren aktuelle Aktivitäten. Hieraus ableitbar sind Aussagen zu möglichen Auswirkungen auf deren Gesundheit und Wohlbefinden.

Im Folgenden wurde in der eigenen Arbeit (Roletscheck, Watzka, [Seiderer](#), Schiller & André, 2018) eine automatische Erkennung von Umgebungsgeräuschen untersucht. Die Evaluation fand im Rahmen der DCASE³⁴ Challenge 2018 statt, bei der von verschiedenen Personen / Gruppen unterschiedlichste Erkennungsmethoden verglichen werden konnten. Tiefe neuronale Netze liefern derzeit die besten Ergebnisse, allerdings stellt die möglichst optimale Bestimmung der Architektur und Hyperparameter eine Herausforderung dar. Die hier vorgestellte Lösung nutzt einen neuartigen, genetischen bzw. evolutionären Algorithmus namens „DeepSAGA“ (Deep Self-Adaptive-Genetic-Algorithm), der zur automatischen Ermittlung der Architektur und Hyperparameter für möglichst optimale tiefe neuronale Netze eingesetzt wurde. Die somit erzeugten neuronalen Netze wurden mit dem Datensatz für Geräuschumgebungen (acoustic scenes) evaluiert.

Verwandte Arbeiten

Die Entwicklung der Architektur von neuronalen Netzen und das Finden zugehöriger Hyperparameter stellt heutzutage, selbst für Experten, eine komplexe und aufwändige Aufgabe dar. Es existieren zwar grundsätzliche Regeln für potenziell gut geeignete Netzstrukturen für bestimmte Probleme, diese grenzen die Anzahl an Möglichkeiten jedoch nur sehr eingeschränkt ein. Bereits seit einiger Zeit spielt daher die Automatisierung dieses Vorgangs in der Forschung eine bedeutende Rolle, wie beispielsweise folgende Arbeiten zeigen: (Fernando et al., 2016; Jozefowicz et al.,

³Detection and Classification of Acoustic Scenes and Events

⁴<http://dcase.community/challenge2018/task-acoustic-scene-classification> (25.09.2019)

2015; M. Kim & Rigazio, 2015; Morse & Stanley, 2016). Neben anderen Verfahren sind evolutionäre Algorithmen eine häufig verwendete Option, um neuronale Netzwerke automatisch zu generieren.

Eine der ältesten Arbeiten bezüglich der Generierung von neuronalen Netzen ist von Miller (1989). In dieser wurde allerdings nicht die gesamte Architektur erzeugt, sondern es wurden nur die Gewichtungen weiterentwickelt. Erst in der Arbeit von Stanley und Miikkulainen (2002) wurde ein Verfahren namens „NeuroEvolution of Augmenting Topologies“ (NEAT) vorgestellt, das Architektur und Gewichtungen modifiziert und bis heute eine Grundlage für viele Algorithmen bildet.

Eine angepasste Version des NEAT-Algorithmus wurde von Kroos und Plumbley (2017) vorgestellt. Er erzeugt kleine neuronale Netze zur Erkennung von Geräuschereignissen (sound events). Die kleinen Netze schnitten in der DCASE-Challenge 2017 im Vergleich zu größeren Netzen ähnlich gut ab. Der genutzte Ansatz zielte allerdings auf eine möglichst geringe Anzahl an Neuronen bzw. Knoten ab und nicht darauf, die Leistung der Klassifizierung zu maximieren.

In der Arbeit von Real et al. (2017) wurden die Möglichkeiten von evolutionären Algorithmen bei der Erzeugung von „Convolutional Neural Networks“ (CNNs) zur Bilderkennung aufgezeigt. Die hierdurch generierten Netze konnten mit den besten manuell erstellten Architekturen mithalten. Der Nachteil der eingesetzten Technik lag darin, dass eine hohe Rechenleistung bei der Suche nach den besten Modellen benötigt wurde. Eine Rechenzeit von 256 Stunden, verteilt auf 250 Clients, wurde benötigt, um das beste Modell zu finden. Durch diesen hohen Rechenbedarf ist das Verfahren in vielen Fällen nicht praktikabel einsetzbar, es zeigte allerdings, wie gut evolutionäre Algorithmen dazu geeignet sind, neue Architekturen zu finden.

Martín et al. (2018) entwickelten einen neuartigen evolutionären Algorithmus, der die Parameter und die Architektur der neuronalen Netze weiterentwickelt, um die Klassifikationsgenauigkeit zu maximieren. Es wird hierbei stets dafür gesorgt, dass valide Netzstrukturen erzeugt werden. Die Parameter des evolutionären Algorithmus wurden hier allerdings per Hand ausgewählt und während der Laufzeit nicht weiter angepasst.

Die verschiedenen vorgestellten Verfahren zeigen die generellen Möglichkeiten evolutionärer Algorithmen zur Generierung neuronaler Netze und dem Auffinden möglichst optimaler Hyperparameter. Der DeepSAGA-Algorithmus wurde entwickelt, da die bereits existierenden Lösungen verschiedene Einschränkungen besaßen. Das Ziel von DeepSAGA bestand darin, eine möglichst optimale Lösung zu finden, wobei eine beschränkte Rechenleistung und Zeitdauer beachtet werden mussten. Wie auch einige andere Algorithmen entwickelt DeepSAGA die Netzarchitektur sowie die Hyperparameter weiter und verwendet dabei den Backpropagation-Algorithmus (Rumelhart et al., 1986) zur Optimierung der Kantengewichte. Eine Besonderheit war, dass die Parameter des evolutionären Algorithmus ebenfalls in der evolutionären Suche enthalten waren (self-adaptive) und diese somit nicht zuvor bestimmt werden mussten.

DeepSAGA-Algorithmus

Für die Entwicklung des Algorithmus wurden die Leitlinien zum Design von ausführbaren evolutionären Algorithmen aus dem Buch von Eiben und Smith (2015) berücksichtigt. Details können den angegebenen Kapiteln entnommen werden.

Eine der Leitlinien schlägt die Verwendung einer Parameterkontrolle vor, da diese das Auffinden passender Parameter vereinfacht (Eiben & Smith, 2015, Kapitel 7.3). Es wurde in diesem Fall eine selbstanpassende (engl. self-adaptive) Variante der Parameterkontrolle eingesetzt, bei der die Parameter des evolutionären Algorithmus ein Teil des Genoms sind und sich somit im evolutionären Suchraum befinden. Der Algorithmus passt sich daher zur Laufzeit an das Problem an, während er gleichzeitig versucht, es zu lösen (Eiben & Smith, 2015, Kapitel 8).

Die Repräsentation und Definition der Individuen, die verwendete Fitnessfunktion und Details zur Population, werden im Nachfolgenden beschrieben. Der Begriff *Session* beschreibt in diesem Abschnitt den vollständigen Ablauf des evolutionären Algorithmus, von der Initialisierung bis zur

Terminierung. Als *Zyklus* werden die sich wiederholenden Schritte der Elternselektion, Rekombination, Mutation, Evaluation der Kinder und die Selektion von Individuen für die nächste Population bezeichnet. Der Ablauf ist in Abbildung 7.2 zusammengefasst.

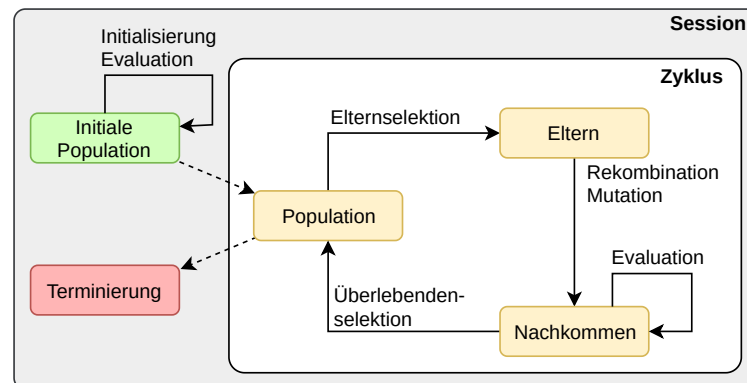


ABBILDUNG 7.2: Genereller Ablauf eines evolutionären Algorithmus.

Repräsentation und Definition Die Repräsentationen werden in dieser Arbeit analog zur biologischen Genetik verwendet, da die Grundidee evolutionärer Algorithmen aus der Biologie stammt (Eiben & Smith, 2015, Kapitel 2.3.2). In der biologischen Terminologie werden phänotypische Eigenschaften, das Äußere eines Lebewesens, von einer genotypischen, inneren Ebene repräsentiert. Im **Genotyp** sind alle Informationen enthalten, um einen bestimmten **Phänotyp** zu erstellen. Das **Genom** enthält alle Chromosomen und repräsentiert die vollständige Genetik eines Lebewesens. Ein **Chromosom** besteht aus Genen, wobei die **Gene** die verschiedenen Eigenschaften des Organismus bestimmen. Ein **Allel** stellt eine konkrete Ausprägung eines Gens dar. Beispielsweise wäre bei einer Blütenpflanze die Blütenfarbe ein Gen. Für konkrete Farben wie Gelb oder Rot existieren bestimmte Allele. Es handelt sich hierbei allerdings um eine Vereinfachung, da ein bestimmtes Gen in der Natur auch mehrere äußere Eigenschaften beeinflussen kann (Eiben & Smith, 2015, Kapitel 2.3.2). Dies wird allerdings in dieser und vielen anderen technischen Umsetzungen nicht nachgebildet.

Beim evolutionären Algorithmus repräsentiert das Genom die Genetik des neuronalen Netzes und beschreibt dessen Charakteristik, die durch die Architektur und die Hyperparameter festgelegt wird. Der Genotyp wird durch das Genom und die Dekodierung ausgedrückt. In diesem Fall beschreibt der Genotyp den Prozess der Erstellung und des Trainierens eines neuronalen Netzes. Ein Phänotyp stellt eine Lösung für einen bestimmten Genotyp dar.

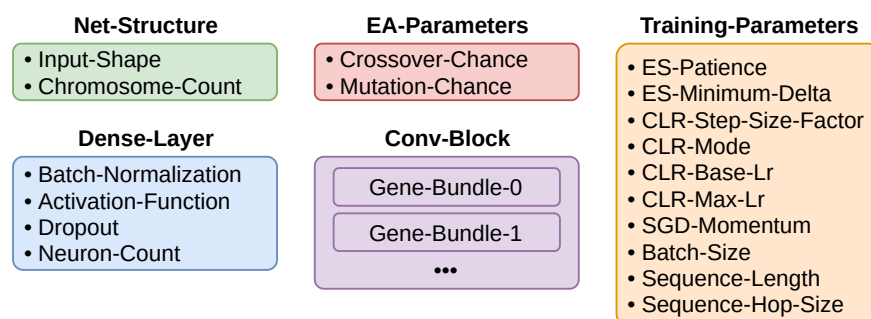


ABBILDUNG 7.3: Überblick über alle Chromosomen und Gene eines Genoms. ES: early stopping, CLR: cyclic-learning-rate.

Abbildung 7.3 zeigt die Liste der Chromosomen in jedem Genom. Die Farben in den folgenden Abbildungen sind zusammengehörig. Der *Conv-Block* besteht aus mindestens einem *Gene-Bundle*.

Für jedes *Gene-Bundle* wird eine Convolutional-Schicht, gefolgt von einer Max-Pooling-Schicht, zur Architektur eines neuronalen Netzes hinzugefügt. Das *Gene-Bundle* enthält daher Informationen (Gene) über die Anzahl der Filter, die Filter-Form (engl. shape) und den Filter-Schritt (engl. stride). Zusätzlich enthält jedes *Gene-Bundle* Gene, die angeben, ob optionale Schichten für Zero-Padding, Dropout und Batch Normalization enthalten sind. Eine Global-Average-Pooling- oder Flatten-Schicht kann verwendet werden, um den *Conv-Block* mit der Ausgabeschicht oder weiteren Schichten zu verbinden. Ein Beispiel für ein Allel für das numerische *Batch-Size-Gen* wäre die Integer-Zahl „128“. Abbildung 7.4 zeigt die Struktur des gesamten Genoms.



ABBILDUNG 7.4: Aufbau eines Genoms. Die Chromosome sind abgebildet: Net-Structure, Conv-Block, Dense-Layer, Training-Parameters und EA-Parameters.

Fitnessfunktion Bei dieser Arbeit lag das Hauptaugenmerk darin, ein neuronales Netz mit möglichst hoher Genauigkeit zu erhalten. Der evolutionäre Suchvorgang sollte allerdings beschleunigt werden, sodass auch die Anzahl der Trainingsepochen berücksichtigt wurde. Der *Score*-Wert gibt die Gesamtfitness eines Individuums in der Population an. Desto höher der Score, desto höher ist die Qualität eines Genotyps. Formel 7.1 entspricht der verwendeten Fitnessfunktion.

$$\text{score} = 0.98 \cdot \text{accuracy} + 0.02 \cdot \frac{\text{epoch}_{\text{limit}} - \text{epoch}}{\text{epoch}_{\text{limit}}} \quad (7.1)$$

Die Maximalzahl der erlaubten Epochen wird durch die Variable $\text{epoch}_{\text{limit}}$ angegeben, die Variable epoch gibt die Epochenzahl an, mit der das neuronale Netz aktuell trainiert wurde. Die Gewichtungen (0.98 und 0.02) basieren auf Beobachtungen und lieferten in Testläufen gute Ergebnisse.

Population Für das Verwalten der Population wurde das Steady-State-Modell verwendet. Im Gegensatz zum Generations-Modell werden beim Steady-State-Modell nur Teile der Population ersetzt (Eiben & Smith, 2015, Kapitel 5.1). Die Größe der Population war dynamisch, um die benötigte Rechenleistung und Laufzeit zu begrenzen, konnte sie allerdings maximal 90 Individuen erreichen.

Elternselektion Wie von Bäck et al. (2000) beschrieben, wird eine Selektion der Eltern mittels Turnier (engl. tournament) durchgeführt. Die Anzahl der erlaubten Teilnehmer wird mit Formel 7.2 bestimmt. Es werden die aktuelle ($\text{popsize}_{\text{current}}$) und die maximal erlaubte Größe der Population ($\text{popsize}_{\text{limit}}$) berücksichtigt.

$$\text{participants} = \text{popsize}_{\text{limit}} - \text{popsize}_{\text{current}} \quad (7.2)$$

Die Größe des Turniers wird mit Formel 7.3 berechnet. Die Variable $\text{toursize}_{\text{limit}}$ ist hierbei immer ein Zehntel von $\text{popsize}_{\text{limit}}$.

$$\text{toursize} = \text{toursize}_{\text{limit}} \cdot \frac{\text{popsize}_{\text{current}}}{\text{popsize}_{\text{limit}}} \quad (7.3)$$

Falls das Populations-Limit erreicht wird, wird die Teilnehmerzahl auf zwei begrenzt, bis der Schwellwert wieder unterschritten wird.

Variationsoperatoren Bei der eingesetzten Sampling-**Mutation** werden Gene der Kategorie „Symbolisch“ (Datentyp: String) in der Art mutiert, dass Allele zufällig mit einem anderen ersetzt werden. Es kann jedoch auch ein identisches Allel ausgewählt werden, wodurch es zu keiner Änderung kommt. Die Werte, die ein Allel annehmen kann, werden durch eine Menge vorgegeben.

Ein numerisches Allel vom Typ „Integer“ wird durch eine Creep-Mutation angepasst oder zurückgesetzt (Wahrscheinlichkeit: 5 %). Bei einer Creep-Mutation wird ein kleiner positiver oder negativer Wert mit einer bestimmten Wahrscheinlichkeit addiert (Eiben & Smith, 2015, Kapitel 4.3.1). Der Sigma-Wert liegt hier immer beim 0,025-fachen des maximal erlaubten Limits.

Ein numerisches Allel vom Typ „Float“ wird mittels einer Nonuniform-Mutation verändert. Eine Nonuniform-Mutation verhält sich analog zur Creep-Mutation für Integer-Werte (Eiben & Smith, 2015, Kapitel 4.4.1). In diesem Fall wird der Sigma-Wert mittels der Mutationswahrscheinlichkeit des Individuums berechnet.

Jeder Teilnehmer der Population hat eine eigene Mutationswahrscheinlichkeit, die in der Form coevolviert wird, wie es in der Arbeit von Bäck (1992) beschrieben wurde. Vor allen anderen Genen wird das *Mutation-Chance*-Gen mit der Nonuniform-Mutation mutiert. Die dadurch ermittelte neue Mutationswahrscheinlichkeit wird für die Mutation der übrigen Gene verwendet.

Zur **Rekombination** wurde das Crossover-Verfahren verwendet. Die Wahrscheinlichkeit, mit der diese stattfindet, hängt von der Crossover-Wahrscheinlichkeit ab. Diese wird durch das *Crossover-Chance*-Gen repräsentiert, wodurch jeder Teilnehmer der Population eine individuelle Wahrscheinlichkeit besitzt. Der Vorgang der Rekombination basiert auf dem Verfahren, das in (Eiben & Smith, 2015, Kapitel 8.4.7) beschrieben wird. Die Crossover-Wahrscheinlichkeit eines potenziellen Elternteils p_c wird mit einer Zufallszahl r verglichen, wobei $r \in [0, 1]$. Ein Elternteil ist bereit sich zu paaren, wenn $p_c > r$. Hierdurch können folgende Fälle auftreten:

1. Beide Elternteile sind bereit sich zu paaren → ein Crossover findet statt.
2. Beide Elternteile sind **nicht** bereit sich zu paaren → sie werden geklont.
3. Nur ein Elternteil ist bereit sich zu paaren → ein Klon des nicht bereiten Elternteils wird erstellt und für den anderen Elternteil zufällig ein neuer Partner aus dem Elternpool ausgewählt, solange bis ein paarungsbereiter Partner gefunden wird.

Die Rekombination findet mittels eines Uniform-Crossovers (Eiben & Smith, 2015, Kapitel 4.2.2) statt. Zum Beispiel erhält Nachkomme A zunächst alle Chromosomen von Elternteil A. Daraufhin wird für jedes Chromosom des Nachkommens die Crossover-Wahrscheinlichkeit von Elternteil A berücksichtigt, um zu entscheiden, ob ein Chromosom mit dem von Elternteil B ausgetauscht wird. Siehe hierzu auch Abbildung 7.5.

Selektion von Überlebenden Das Selektionsverfahren folgt einer altersabhängigen Ersetzungsstrategie (siehe hierzu Eiben und Smith, 2015, Kapitel 5.3.1). Jedem neu erstellten Individuum wird ein Wert als verbleibende Lebensdauer (RLT) zugewiesen. Es wird hierzu Formel 7.4, wie von Bäck et al. (2000) beschrieben, verwendet. Nach jedem Zyklus wird RLT bei allen Individuen, außer dem Individuum mit der höchsten Fitness, um eins verringert. *MinLT* und *MaxLT* stehen für die zulässige minimale und maximale Lebensdauer. Die anderen Variablen beziehen sich auf den aktuellen Status der Population: die Fitness des Individuums i ($fitness(i)$), die mittlere (*AvgFit*), beste (*BestFit*) und die schlechteste Fitness (*WorstFit*). Die Berechnung des Vorfaktors wird mit folgender Formel durchgeführt: $\eta = \frac{1}{2} \cdot (MaxLT - MinLT)$.

$$RLT(i) = \begin{cases} MinLT + \eta \cdot \frac{WorstFit - fitness(i)}{WorstFit - AvgFit} & \text{if } fitness(i) \geq AvgFit \\ \frac{1}{2} (MinLT + MaxLT) + \eta \cdot \frac{AvgFit - fitness(i)}{AvgFit - BestFit} & \text{if } fitness(i) < AvgFit \end{cases} \quad (7.4)$$

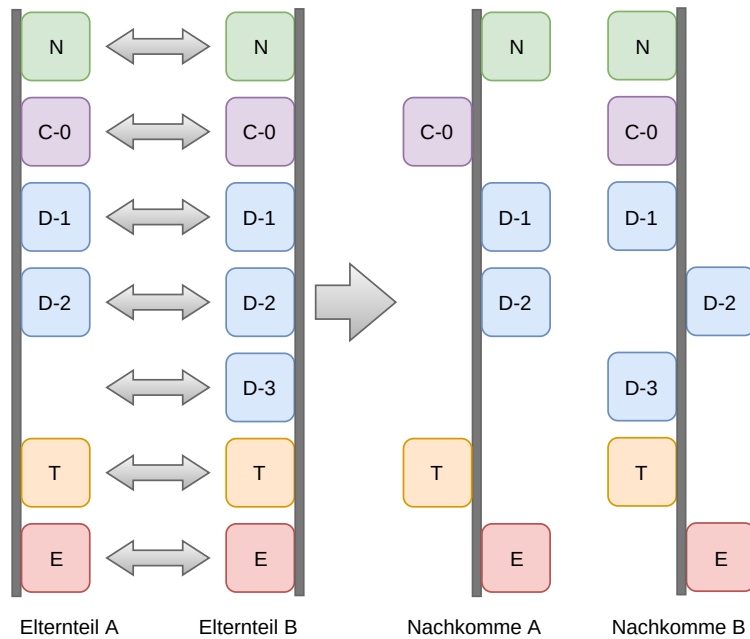


ABBILDUNG 7.5: Beispiel zur Rekombination durch Uniform-Crossover.

Die minimale und maximale Lebensdauer wurden in den Versuchen auf 1 bzw. 7 gesetzt. Falls der Fitnesswert eines neuen Individuums besser als die mittlere Fitness ist, wird eine Lebensdauer von 5 und 7 zugewiesen, ansonsten 1 und 4. Hierdurch besitzen bessere Individuen eine längere Lebensdauer als schlechtere.

Initialisierung und Terminierung Die Population der ersten Generation wird zufällig erzeugt. Allele können nur Werte mit festgelegten Einschränkungen zugewiesen bekommen, die auch während der Evolution berücksichtigt werden, um die Komplexität zu reduzieren. Terminiert wird der Evolutionäre-Prozess nach dem 40. Zyklus, da das Optimum vorher nicht bekannt ist.

Datensatz

Der Datensatz, der für die DCASE-Challenge 2018 Acoustic Scene Classification (Subtask A) von den Organisatoren bereitgestellt wurde, war der TUT-Urban-Acoustic-Scenes-2018-Datensatz (Mesaros et al., 2018). Er besteht aus folgenden zehn Klassen:

- Flughafen (airport)
- innerhalb Einkaufszentrum (indoor shopping mall)
- Metro-Station (metro station)
- Fußgängerstraße (pedestrian street)
- öffentlicher Platz (public square)
- mittelstark-befahrene Straße (street with medium level of traffic)
- Straßenbahnfahrt (travelling by a tram)
- Busfahrt (travelling by a bus)
- U-Bahnfahrt (travelling by an underground metro)
- städtischer Park (urban park)

Die Daten wurden in sechs europäischen Großstädten an verschiedenen Orten aufgezeichnet. Pro Klasse sind 864 Dateien mit einer Länge von 10 Sekunden (144 Minuten pro Klasse) vorhanden. Der Entwicklungs-Datensatz wurde aufgeteilt, sodass 6122 Dateien von allen Klassen für das Training (ca. 70 %) und 2518 Dateien für den Test-Datensatz zur Verfügung standen. 3600 Dateien wurden für die Evaluation genutzt. Deren Beschriftungen waren allerdings nur den Organisatoren bekannt, da diese zum Vergleich der Verfahren verschiedener Personen / Teams genutzt wurden.

Merkmalsextraktion

Für die eigenen Experimente wurden die Stereo-Dateien in Mono umgewandelt und mittels des „librosa“ (v0.6.1) Python-Moduls (McFee et al., 2015) Log-Mel-Spektrogramme mit 100 Bändern erzeugt, die einen Frequenzbereich von 22 050 Hz umfassten. Für die Kurzzeit-Fouriertransformation (STFT) wurden Fenster mit einer Länge von 2048 Samples (43 ms), einer Überlappung von 50 % mit der Hamming-Fensterfunktion eingesetzt. Die entstandenen Spektrogramme wurden in Sequenzen mit einer bestimmten Sequenzlänge, mit einer Überlappung von 50 %, aufgeteilt. Die Sequenzlänge war abhängig von den durch den genetischen Algorithmus generierten Modellen.

Evaluation

Die initiale Population des genetischen Algorithmus wurde für die Evaluation nicht zufällig erzeugt. Zwei unabhängige Sessions mit jeweils 10 Zyklen wurden zuvor durchgeführt und die jeweils 30 besten Modelle für die initiale Population der Evaluations-Session ausgewählt. Hierdurch konnte bereits am Anfang eine höhere Fitness der Population gewährleistet werden, wobei eine relativ hohe Diversität noch vorhanden war.

Um den Vorgang zu beschleunigen, wurde die Rechenlast mittels Client-Server-Verfahren auf mehrere Rechner verteilt. Der Server verteilt die Genotypen der Population des aktuellen Zyklus an alle Clients, bei denen die Dekodierung durchgeführt wird. Clients schicken nach der Verarbeitung den entstandenen Phänotyp (Lösung) an den Server zurück. Insgesamt waren maximal 15 Rechner mit jeweils einer NVIDIA GTX-1060-Grafikkarte, zur Beschleunigung des Trainingsvorgangs, im Einsatz. Abhängig von der aktuellen Größe der Population, der Komplexität des Genoms und der zur Verfügung stehenden Rechner dauerte ein Zyklus ca. zwei bis drei Stunden. Die insgesamt benötigte Laufzeit lag bei ca. 120 Stunden. Ohne die zwei initialen Sessions betrug die Laufzeit ca. 87 Stunden.

Am Ende der letzten Session wurde das beste neuronale Netz für die Klassifikation verwendet. Zusätzlich wurde die Ensemble-Learning-Strategie „Population vote“ eingesetzt. Hierfür wurden die zehn besten Individuen ebenfalls genutzt, um zusammen für die Klasse eines Audio-Samples zu stimmen. Individuen mit höherer Fitness bekamen mehr Stimmen. Die Klasse mit den meisten Stimmen gewinnt und wird somit als Ergebnis verwendet.

Ergebnisse

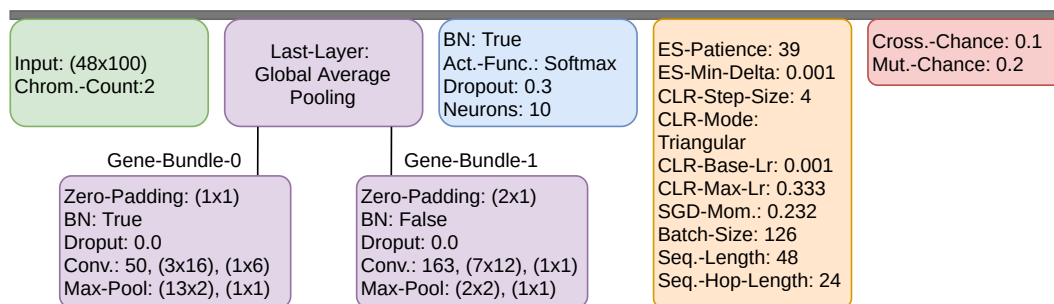
In Tabelle 7.1 sind die finalen Klassifikationsergebnisse dargestellt, die auf dem Test-Datensatz basieren. Das beste Ergebnis (mittlere Treffergenauigkeit von 72,8 %) wurde mit dem „DeepSAGA CNN“ erreicht und übertrifft die Baseline (mittlere Treffergenauigkeit von 59,7 %) deutlich. Durch den Einsatz des Population-Vote-Verfahrens konnte, das Ergebnis noch weiter verbessert werden (mittlere Treffergenauigkeit von 74,7 %).

Beim Evaluations-Datensatz (Beschriftungen nur den Organisatoren bekannt), erreichte das beste Ergebnis eine Treffergenauigkeit von 69,2 %, wobei die Baseline hier bei 61,0 % lag. Die besten zwei Verfahren, die allerdings per Hand entwickelt wurden, erreichten bei der Challenge 81,0 %.

Das Genom des „DeepSAGA CNN“ ist in Abbildung 7.6 dargestellt. Abbildung 7.7 zeigt die generierte Architektur, die auf Basis des Genoms erzeugt wurde. Die verwendeten Hyperparameter sind in Abbildung 7.6 dem *Training-Parameters*-Chromosom zu entnehmen.

Beschriftung	DCASE2018 Baseline	DeepSAGA CNN	Pop. Vote
Airport	72,9 %	84,9 %	85,7 %
Bus	62,9 %	63,2 %	67,4 %
Metro	51,2 %	71,3 %	71,6 %
Metro station	55,4 %	75,3 %	81,9 %
Park	79,1 %	81,0 %	82,2 %
Public square	40,4 %	53,2 %	56,0 %
Shopping mall	49,6 %	75,3 %	73,8 %
Street, pedest.	50,0 %	67,2 %	69,6 %
Street, traffic	80,5 %	85,0 %	86,2 %
Tram	55,1 %	72,0 %	72,4 %
Mittelwert	59,7 %	72,8 %	74,7 %

TABELLE 7.1: Treffergenauigkeit der einzelnen Klassen auf dem Test-Datensatz.

ABBILDUNG 7.6: Bestes Genom, der finalen Session. BN: Batch Normalization. Die Zahlen in den Klammern geben die Filtergröße (*filter-size*) und den Filter-Schritt (*filter-stride*) bei den Convolutional- und Max-Pooling-Schichten an. Die erste Zahl bei Convolutional-Schichten steht für die Filteranzahl.

Diskussion und Zusammenfassung

In dieser Untersuchung wurde ein genetischer Algorithmus namens „DeepSAGA“ entwickelt, der automatisch Convolutional Neural Networks (CNNs) erzeugen kann. Sobald der Vorgang gestartet wird, arbeitet der Algorithmus selbstständig und stellt am Ende eine Auswahl an neuronalen Netzen zur Verfügung. Die DCASE 2018 Challenge diente als Anwendungsszenario und sollte zeigen, wie sich das automatische Verfahren im Vergleich zu anderen Systemen schlägt, die zum Großteil per Hand erstellt wurden. Die Resultate waren hierbei vielversprechend.

Die Beobachtungen während einer gesamten Session zeigten, dass die Einbeziehung der Hyperparameter im Suchvorgang eine bedeutsame Entscheidung war. Einige Klone unterschieden sich ausschließlich dadurch, dass einzelne Hyperparameter (*Training-Parameter-Chromosom*) verschieden war. Bereits diese Änderungen verursachten gravierende Unterschiede bezüglich ihrer Treffergenauigkeiten. Daraus lässt sich schließen, dass bestimmte Architekturen nur mit bestimmten Hyperparametern gut funktionieren, was auch umgekehrt der Fall ist.

Alle generierten CNNs zeigten bei der Klasse „public square“ die schlechteste Leistung. Dieses Problem trat auch bei der Baseline auf. Es zeigte sich, dass diese Klasse häufig mit „shopping mall“ oder „street traffic“ verwechselt wurde. Bei all diesen Klassen waren im Hintergrund Gespräche hörbar und ausgenommen von der „shopping mall“-Klasse waren zusätzlich Verkehrsgeräusche enthalten. In zukünftigen Untersuchungen könnten daher Lösungsansätze getestet werden, wie dieses Problem beim genetischen Algorithmus berücksichtigt werden könnte.

Der entwickelte Algorithmus ist bei der Erstellung der Architektur folgendermaßen eingeschränkt: nach der Eingabeschicht können nur eine Reihe von Convolutional-Schichten gefolgt von Dense-Schichten folgen. Hierdurch kann zum Beispiel auf eine Convolutional-Schicht, gefolgt von einer Dense-Schicht keine weitere Convolutional-Schicht folgen. Zudem wurden keine

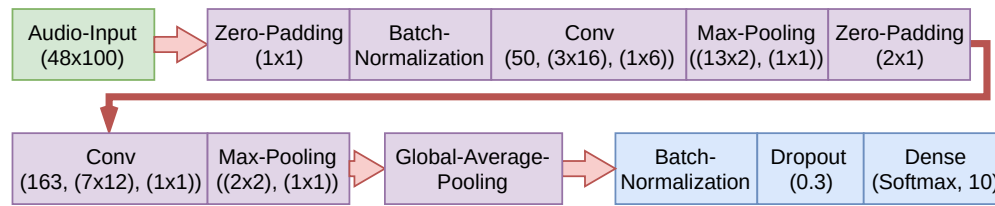


ABBILDUNG 7.7: Architektur des besten Netzes: „DeepSAGA CNN“.

Recurrent-Schichten einbezogen, die gerade bei kontinuierlichen Daten Vorteile besitzen. Diese beiden Ergänzungen könnten zu deutlich besseren Erkennungsraten führen, allerdings müsste eine Vielzahl neuer Parameter im evolutionären Prozess berücksichtigt werden. Dadurch würde die Laufzeit wesentlich erhöht und müsste anderweitig reduziert werden.

Der DeepSAGA-Algorithmus ist nicht auf Audiodaten beschränkt. Allerdings wären weitere Untersuchungen nötig, um die Nutzbarkeit mit verschiedenen Datenarten und Datensätzen zu prüfen und eventuell notwendige Optimierungen am Algorithmus durchzuführen. Der Algorithmus stellte allerdings einen weiteren Schritt dar, das aufwändige, manuelle Erstellen von neuronalen Netzen unnötig zu machen oder den Menschen hierbei zumindest zu unterstützen.

7.3.2 Auswirkungen der städtischen Umgebung auf den Menschen

Einleitung

Durch den Klimawandel treten immer höhere Temperaturen vor allem auch in Städten auf. Daher werden bei der Stadtplanung verschiedene Ansätze untersucht, um die Wärmebelastung beim Menschen zu reduzieren. Hierzu zählen beispielsweise Schwammstädte (engl. *Sponge-Cities*), bei denen u. a. mit Bepflanzung verhindert wird, dass immer häufiger auftretende Niederschläge zu Überschwemmungen führen (Chan et al., 2018). Die Bepflanzung speichert hierbei eine große Menge an Wasser und gibt es anschließend langsam durch Verdunstung und Fotosynthese ab, wodurch zudem die Erhitzung der Umgebung verlangsamt wird. Dies wird bei größeren Büschen oder Bäumen zusätzlich dadurch ermöglicht, dass sie Schatten auf Gebäude oder Straßen werfen. Die Begrünung von Städten hat daher das Potenzial, den vom Klima verursachten Stress bei Personen zu reduzieren. Bepflanzung kann jedoch auch andere positive Auswirkungen besitzen. Naturtherapien, wie Gartentherapie (Adevi & Mårtensson, 2013) oder Waldbaden (Park et al., 2010), das auch als *Shinrin-yoku* (japanisch) bezeichnet wird, zeigten positive Auswirkungen auf den menschlichen Verstand und Körper (Kotera et al., 2020; Rathmann & Brumann, 2017). Bereits der Blick durch ein Fenster auf natürliche Umgebung könnte laut Ulrich (1984) einen gewissen gesundheitsfördernden Effekt besitzen.

Bezüglich der Mensch-Computer-Interaktion können in der Literatur mehrere Arbeiten gefunden werden, die mit Technologie die Interaktion zwischen Mensch und Natur verbessern möchten. Hierzu zählt beispielsweise die Interaktion mit Zimmerpflanzen (Bittner et al., 2019). Einige der Arbeiten versuchten zudem, die Ästhetik von Landschaften zu berücksichtigen, um das Wohlbefinden der Nutzer zu verbessern. Hierzu zählt z. B. die Arbeit von Runge et al. (2016), in der ein Navigationssystem präsentiert wurde, das die Route mit der schönsten Szenerie vorschlagen konnte. Vergleichbar ist das Empfehlungssystem für Bustouren von Samsonov et al. (2017), bei dem der Nutzer eine Empfehlung für die Seite vom Bus erhielt, auf der die Sicht auf die Landschaft aus dem Fenster am schönsten war. Hierbei wurden allerdings keine objektiven Messdaten hinsichtlich einer Wohlbefindensveränderung der Busreisenden erfasst. In der Literatur finden sich zur Bestimmung des Wohlbefindens verschiedene Modelle. Neben subjektiven Metriken können auch auf physiologischen Daten basierende objektive Messwerte bestimmt werden (Kalimeri & Saitis, 2016; K. H. Kim et al., 2004). Die direkten Einflüsse der Umwelt auf die Physiologie wurde

durch die Messung der Herzrate und deren Variabilität in der Arbeit von Song et al. (2014) untersucht. Nakayoshi et al. (2015) analysierten die Veränderungen der Hauttemperatur und Herzrate abhängig von der Umgebung.

Ziel der eigenen Untersuchung war es zu ermitteln, inwiefern das Wohlbefinden eines Menschen von der städtischen Umgebung und deren lokalen Klimazonen beeinflusst wird. Hierfür wurden physiologische Daten, Umweltdaten und subjektive Bewertungen des Wohlbefindens mit mobilen Geräten aufgezeichnet und anschließend ausgewertet. Die Auswertung erfolgte mittels maschinellem Lernen. Die hierbei verwendeten Techniken wurden so ausgewählt, dass sie in Echtzeit auch auf einem mobilen Gerät zur Klassifizierung genutzt werden könnten.

Die Untersuchung wurde in den Publikationen (Flutura, [Seiderer](#), Aslan, Dietz, Schiller, Beck, Rathmann & André, 2019) und (Rathmann, Beck, Flutura, [Seiderer](#), Aslan & André, 2020) veröffentlicht. Es handelte sich hierbei um eine interdisziplinäre Kooperation zwischen der Informatik und Geografie an der Universität Augsburg.

Aufzeichnung der Daten

Es wurden die Daten von 7 Personen (2 Frauen, 5 Männer, mittleres Alter: 25 Jahre) in 20 Sitzungen aufgezeichnet und für die Evaluation genutzt. Jede Sitzung dauerte ca. 80 Minuten, in der ein festgelegter Rundweg von ca. 5 km, in Gehgeschwindigkeit, durchgeführt wurde. Die Aufzeichnungen fanden im Hochsommer statt und wurden bei ähnlichen Wetterbedingungen und Tageszeiten durchgeführt: wolkenloser Himmel, Mittag bis früher Nachmittag. Es sollten maximale Unterschiede zwischen den verschiedenen lokalen Klimazonen auf der Route auftreten, um potenziellen Hitzestress aufzeichnen zu können.

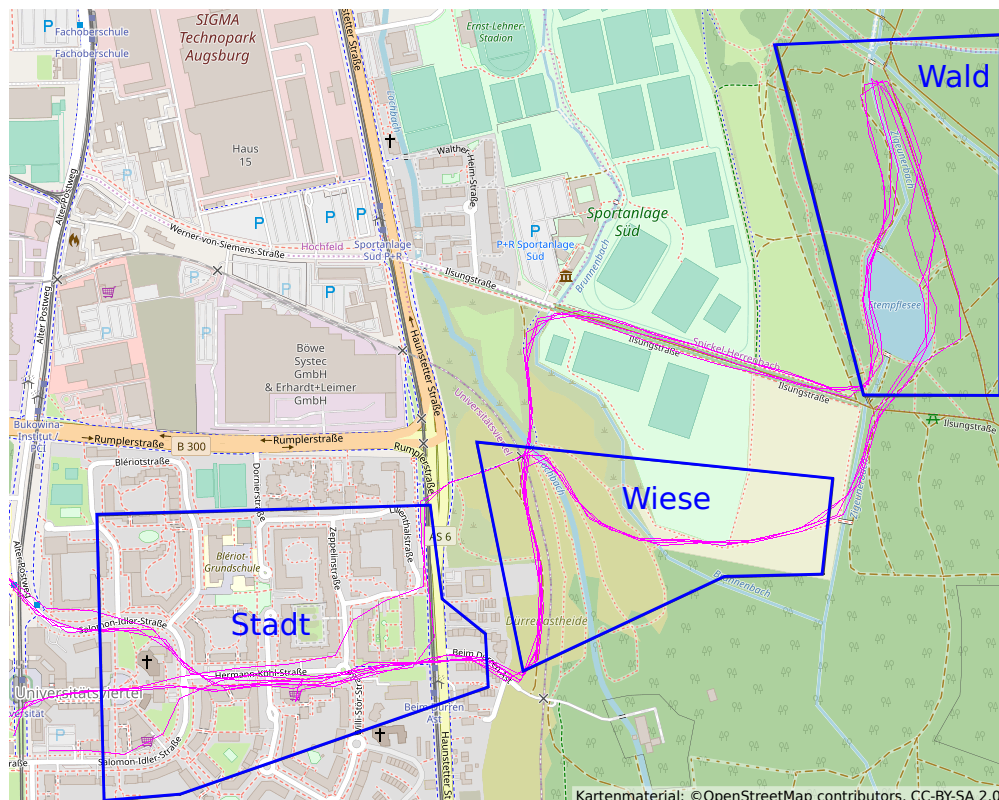


ABBILDUNG 7.8: Lokale klimatische Zonen der gewählten Rundwege (vier Beispiele sind violett eingezeichnet) von der Universität Augsburg aus in den naheliegenden Stadtwald (Siebentischwald). Die entsprechenden Klimazonen (Klassen) entlang der Routen wurden mit Polygonen (blau) festgelegt und für die Evaluation genutzt, wobei Übergänge ausgespart wurden: „Open Mid Rise“ (Stadt), „Scattered Trees“ (Wiese), „Dense Trees“ (Wald).

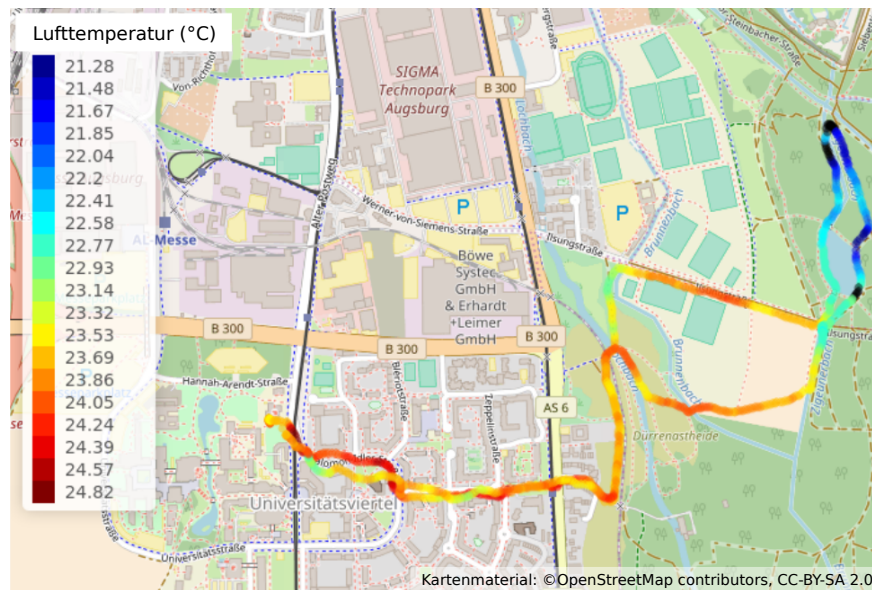


ABBILDUNG 7.9: Lufttemperatur auf der Route bei einer exemplarischen Aufzeichnung. (Bild abgeändert aus (Flutura, **Seiderer**, Aslan, Dietz, Schiller, Beck, Rathmann & André, 2019))

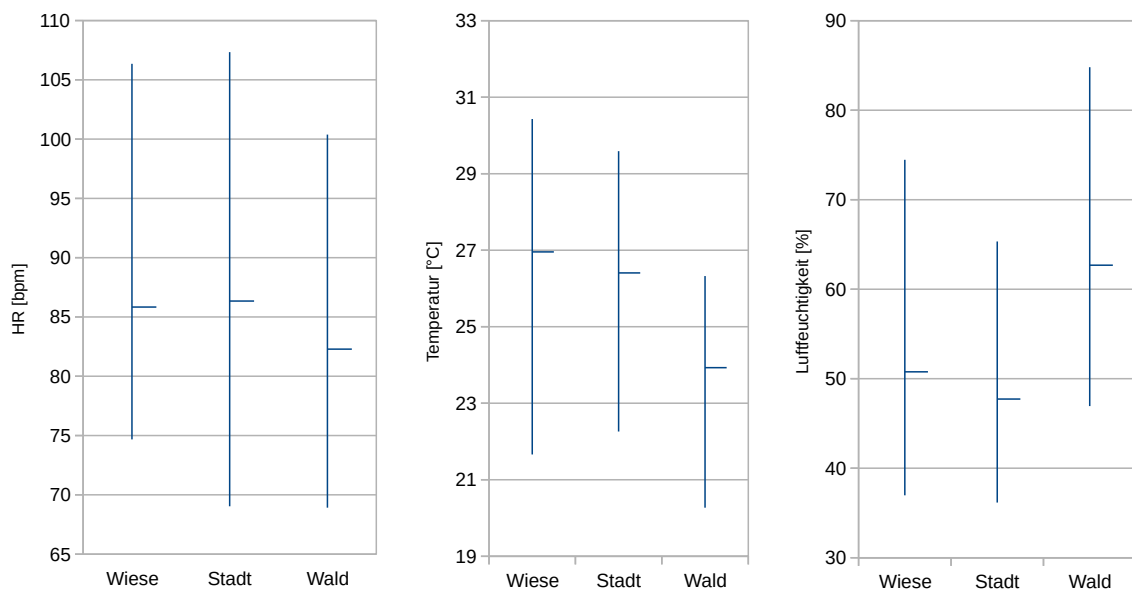


ABBILDUNG 7.10: Über vier Sitzungen (siehe Abbildung 7.8) berechnete minimale, maximale und durchschnittliche Herzrate, Temperatur und Luftfeuchtigkeit in den drei lokalen Klimazonen („Open Mid Rise“ (Stadt), „Scattered Trees“ (Wiese), „Dense Trees“ (Wald)). Die Temperatur- und Luftfeuchtigkeitsdaten stammten hierbei von der am Rucksack befestigten mobilen Gas-Sensorbox (SHT75-Sensor).

Die Route lag im Universitätsviertel der Stadt Augsburg und dem angrenzenden Augsburger Stadtwald (Siebentischwald). Sie ist in Abbildung 7.8 dargestellt. Die Route umfasste die lokalen Klimazonen „Open Mid Rise“, „Scattered Trees“ und „Dense Trees“ (Beck et al., 2018; Stewart & Oke, 2012). Diese lokalen Klimazonen deckten die Wärmebelastung in der Stadt, auf einer offenen Wiese sowie im schattenspendenden Wald, ab. In Abbildung 7.9 sind an einem Beispiel die verschiedenen Temperaturen auf dem Rundweg erkennbar. In Abbildung 7.10 sind die minimale, maximale und durchschnittliche Herzrate, Temperatur und Luftfeuchtigkeit in den drei lokalen Klimazonen über vier Sitzungen dargestellt. Es zeigt sich hierbei, dass sich insbesondere die Messwerte im Wald („Dense Trees“) bei allen drei Diagrammen deutlich abheben.

Gerät	Sensor	Daten	SR (Hz)
Fitnessarmband	BVP-Sensor	IBI (Inter-Beat-Interval)	30
		HR (Heart Rate)	1
	EDA-Sensor	EDA (Elektrodermale Aktivität)	5
	Accelerometer	Beschleunigungsdaten (3-Achsen)	62,5
Smartphone	GPS	Koordinaten	5
	Mikrofon	Audio (raw)	16000
	GUI	Bewertungen	eventbasiert
Smartwatch	GUI	Bewertungen	eventbasiert
Gas-Sensorbox	SDS011	PM2.5	0,07
		PM10	0,07
	SHT75	Luftfeuchtigkeit	0,07
		Temperatur	0,07
	MiCS-6814	CO (Kohlenmonoxid)	0,07
		NO ₂ (Stickstoffdioxid)	0,07
		NH ₃ (Ammoniak)	0,07
		C ₃ H ₈ (Propan)	0,07
		C ₄ H ₁₀ (Butan)	0,07
		CH ₄ (Methan)	0,07
		H ₂ (Wasserstoff)	0,07
		C ₂ H ₅ OH (Ethanol)	0,07
	BMP280	Luftdruck	0,07
		Temperatur	0,07
Aspirations-Psychrometer		Luftfeuchtigkeit	0,5
		GPS-Koordinaten	0,5
		Temperatur	0,5

TABELLE 7.2: Aufgezeichnete Sensordaten der verschiedenen Geräte.

Die Geräte und deren Sensordaten sowie Abtastraten, die zur Aufzeichnung verwendet wurden, sind in Tabelle 7.2 aufgelistet. Ein Überblick über die Kommunikation der Geräte ist in Abbildung 7.11 links dargestellt.

Abbildung 7.11 zeigt rechts die Positionierung der Sensoren. Alle Daten wurden synchronisiert mit einer Kombination aus mobileSSI (siehe Abschnitt 4.1.2) und SSJ auf dem Smartphone (Samsung Galaxy S3) aufgezeichnet. In jeder Aufnahmesitzung wurden die Daten von eins bis zwei Teilnehmern erfasst, wobei ein Teilnehmer die Smartwatch (Samsung Gear S2), der andere das Smartphone zur Annotation mittels kurzer Fragebögen zum aktuellen Wohlbefinden auf dem Display verwendete. Zusätzlich trugen beide Personen jeweils ein Microsoft Band 2 (Fitnessarmband) für die Aufzeichnung von physiologischen Daten. Beim Microsoft Band 2 war neben der weitverbreiteten Erfassung der Beschleunigungsdaten und Herzschläge mittels eines optischen BVP-Sensors ein Hautleitwert-Sensor enthalten. Zudem konnten die Daten in Echtzeit vom Armband mit eigener Software ausgelesen werden. Zur Erfassung der Umweltdaten wurden ein tragbares Aspirationspsychrometer (ALMEMO FNAD 46 mit Datenlogger ALMEMO 2590A) und der eigene mobile Gas-Sensorbox-Prototyp (siehe Abschnitt 3.5.5) mitgeführt.

Ein Aspirationspsychrometer ist ein professionelles Messgerät aus der Meteorologie, das die relative Luftfeuchtigkeit präzise bestimmen kann. Es musste während der Aufzeichnungen an einem Stab in ca. 1,8 - 2,1 m Höhe, gerade nach oben gerichtet, mitgeführt werden. Zudem musste darauf geachtet werden, dass ausreichend destilliertes Wasser im Gerät enthalten war. Das Gerät zeichnete die Temperatur- und Luftfeuchtigkeitswerte autark mit Zeitstempel und GPS-Position auf. Diese Daten konnten daher nachträglich mittels der GPS-Position und dem Zeitstempel synchronisiert werden. Es wurde als Referenz für die Temperatur und Luftfeuchtigkeit genutzt. Für eine praktische Nutzung durch Endanwender wäre es jedoch zu sperrig und hinsichtlich der Handhabung zu kompliziert. Daher war bereits in der mobilen Gas-Sensorbox, die per WLAN

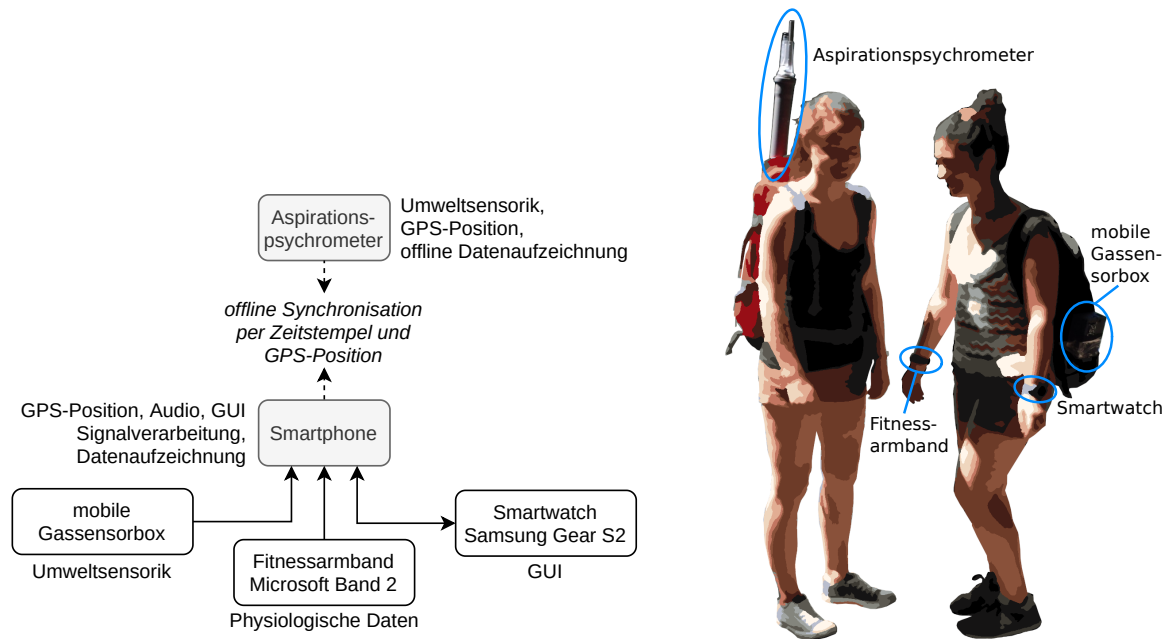


ABBILDUNG 7.11: Links: Übersicht über die Komponenten bei der Datenaufzeichnung. Rechts: Beispiel für Positionierung der Sensoren. (rechtes Bild abgeändert aus (Flutura, [Seiderer](#), Aslan, Dietz, Schiller, Beck, Rathmann & André, 2019))

von mobileSSI ausgelesen werden konnte, ein kompakter, qualitativ hochwertiger Temperatur- und Luftfeuchtigkeitssensor (SHT75) integriert. Der Sensor war allerdings durch das schützende Gehäuse weniger stark exponiert. Das Aspirationspsychrometer und die mobile Gas-Sensorbox waren an jeweils einem Rucksack befestigt, die von den beiden Studienteilnehmern getragen wurden.

Um das momentane subjektive Wohlbefinden der Studienteilnehmer während des Rundgangs zu erfassen, wurden Fragebögen auf den mobilen Geräten eingesetzt, um sogenanntes „Experience Sampling“ durchzuführen. Die Fragen waren bezogen auf die emotionale Valenz der Situation (positiv bis negativ, 5-Punkte-Likert-Skala), wahrgenommene Temperatur (heiß bis kalt, 9-Punkte-Likert-Skala) und Luftqualität (gut bis schlecht, 9-Punkte-Likert-Skala). Laut Frenzel et al. (2009) sind Emotionen ein bedeutsamer Bestandteil des subjektiven Wohlbefindens und können zudem direkte Auswirkung auf die Herzrate und den Hautleitwert besitzen, die in dieser Studie ebenfalls aufgezeichnet wurden. Das Empfinden von Stress, der durch Wärme ausgelöst wurde, sollte durch die Abfrage der wahrgenommenen Temperatur nachvollzogen werden können. Eine erhöhte Umgebungstemperatur besitzt auch objektiv messbare Auswirkungen auf die menschliche Physiologie (z. B. erhöhter Puls und Schweißproduktion). Auch die subjektiv wahrgenommene Luftqualität kann das Wohlbefinden beeinflussen. Bei hohen Konzentrationen von verschiedenen Gasen und Feinstaub können grundsätzlich erhöhte gesundheitliche Risiken auftreten. Sie können jedoch nur bedingt mit mobiler Gassensorik gemessen werden.

Die Studienteilnehmer sollten die Bewertungen immer dann aktualisieren, wenn sie der Meinung waren, dass sich etwas geändert hat. Insgesamt wurden hierbei 769 Bewertungen während der Studie aufgezeichnet. Nach jedem Rundgang wurden zum Abschluss von allen Teilnehmern Fragebögen in Papierform ausgefüllt. Diese beinhalteten ebenfalls Bewertungen bezüglich des Wohlbefindens, der Temperatur und der Luftqualität, die jedoch eine Art Mittelwert darstellten, da sie auf die komplette Route bezogen waren.

Evaluation

Beziehung zwischen lokalen Klimazonen und Physiologie Es konnte eine statistische Signifikanz (nicht-parametrische Kruskal-Wallis-Tests und paarweise Wilcoxon-Tests) zwischen den Temperaturschwankungen entlang des Rundwegs und den lokalen Klimazonen festgestellt werden. Auswirkungen der Umgebung auf die Herzrate konnten ebenfalls direkt beobachtet werden, diese Effekte waren jedoch weniger deutlich zwischen den lokalen Klimazonen unterscheidbar, da hierbei auch andere Faktoren eine Rolle gespielt haben können. Um die Beziehungen zwischen physiologischen Daten, Umgebungsdaten und den Bewertungen der Studienteilnehmer genauer zu analysieren, wurde maschinelles Lernen eingesetzt. Die verwendeten Methoden wurden so ausgewählt, dass sie auch in Echtzeit nutzbar wären, daher wurde beispielsweise versucht möglichst kleine Datenfenster zu nutzen, um zeitnah Ergebnisse bei Klassifikationen erhalten zu können.

Als maschinelles Lernverfahren wurden Support Vector Machines (SVMs) eingesetzt, da diese im Vergleich zu tiefen neuronalen Netzen auch mit kleinen Datenmengen gute Ergebnisse erzielen können. Ein gleitendes Fenster mit einer Größe von 10 Sekunden und einer Überlappung von 240 Sekunden wurde für die Klassifikation der lokalen Klimazone anhand von Herzrate bzw. Intervall zwischen den Herzschlägen (IBI) und Hautleitwert (EDA) eingesetzt. Für die Herzrate und das IBI wurden 22 statistische Merkmale, für den Hautleitwert 64 Merkmale berechnet und anschließend fusioniert (Merkmalsfusion).

Ein hieraus erzeugtes personenabhängiges Klassifizierungsmodell erreichte bei den drei Klassen („Open Mid Rise“, „Scattered Trees“ und „Dense Trees“) insgesamt eine Treffergenauigkeit von 88,33 % (siehe Tabelle 7.3). Es wurden 480 Samples (über die drei Klassen ausbalancierte Sampleanzahl mit „random under-sampling“) mit einer zweifachen Kreuzvalidierung evaluiert. Bei der ausschließlichen Nutzung von EDA-Merkmalen reduzierte sich die Treffergenauigkeit um 14 %. Falls ausschließlich die Herzraten und IBI-Merkmale zur Klassifizierung verwendet wurden, reduzierte sich die Treffergenauigkeit um 36 %.

	Dense Trees	Scattered Trees	Open Mid Rise	Treffergenauigkeit
Dense Trees	129	10	21	80,62 %
Scattered Trees	9	148	3	92,50 %
Open Mid Rise	9	4	147	91,88 %
Mittelwert				88,33 %

TABELLE 7.3: Konfusionsmatrix der personenabhängigen SVM-Klassifizierung der lokalen Klimazonen basierend auf fusionierten Merkmalen von Herzrate, IBI und EDA. 480 Samples aus Fenstern mit einer Größe von 10 Sekunden mit 240 Sekunden Überlappung. Evaluiert mit zweifacher Kreuzvalidierung. Die vorhergesagten Klassen sind die Tabellenspalten.

Daraus lässt sich schlussfolgern, dass die lokalen Klimazonen einen eindeutigen Effekt auf die physiologischen Daten der Studienteilnehmer besaßen. Zudem zeigte das Ergebnis, dass die eingeschränkte Datenqualität eines Fitnessarmbands (Microsoft Band 2) und die Nutzung mobil einsetzbarer Verfahren des maschinellen Lernens, für eine Klassifizierung ausreichend sind. Die Besonderheit des Microsoft Band 2, zusätzlich den Hautleitwert zu erfassen, führte in diesem Fall zu einer deutlichen Verbesserung der Klassifikationsergebnisse.

Beziehung zwischen Umgebung und Wohlbefinden Bereits in der Literatur konnten Beziehungen zwischen natürlicher Klanglandschaft und positiver Emotionen, z. B. in der Analyse von R. T. Buxton et al. (2021), und zwischen physiologischen Signalen und emotionaler Valenz, z. B. in der Arbeit von Vijayakumar et al. (2020), festgestellt werden. Daher wurde zusätzlich untersucht, inwiefern die Nutzerbewertungen der emotionalen Valenz mittels der im Feld aufgezeichneten Sensordaten der mobilen Geräte vorhersagbar waren. Als Klassen des Modells wurden daher die

fünf Bewertungsstufen (negativ bis positiv) des Fragebogens genutzt, der über die Mobilgeräte, während des Laufens, beantwortet wurde.

Zunächst wurden hierfür die aufgezeichneten **Audiodaten** betrachtet, die primär Umgebungsgeräusche beinhalten. In der Stadt waren die Geräusche vor allem vom Menschen erzeugt (anthropogene Quellen), wobei hierzu Straßenverkehr (u. a. PKWs / LKWs, Straßenbahn) und Maschinen zählen. Im Bereich der Wiese und im Wald überwogen natürliche Geräuschquellen wie Vogelgezwitscher (biologische Quelle), wobei im Wald noch zusätzliche Geräusche eines Bachs (geophysische Quelle) sowie von Schwimmvögeln (u. a. Entenvögel) in der Nähe des Sees erfasst werden konnten. Generell wurden zudem kurzfristig Gespräche von Passanten / Studienteilnehmer sowie Geräusche von Wind und Schritten aufgezeichnet. Bereits bei anderen Untersuchungen zeigte sich, dass bei Audioklassifizierungen die Nutzung von tiefen neuronalen Netzen für Bilderkennung auf berechneten Mel-Spektrogrammen gute Ergebnisse erzielen (siehe Abschnitt 5.5.3). Daher wurden diese hierfür ebenfalls eingesetzt. Es wurde die „viridis“-Farbpalette von Matplotlib für die Spektrogramme (224×224 Pixel groß) genutzt und die Merkmalsvektoren des MobileNet_v2_1.4_224 (Netzgröße = 140 %; 224×224 Bildgröße für die Eingabeschicht) mit Tensorflow berechnet. Es handelt sich hierbei um das größte und leistungsfähigste MobileNet v2, das neben anderen in Abschnitt 5.5.2 mit mobilen Geräten bereits getestet wurde. Ein Merkmalsvektor besitzt bei diesem Netz eine Größe von 1792 Werten. Bei einer zehnfachen Kreuzvalidierung konnte eine mittlere, personenabhängige Treffergenauigkeit von 61,79 % über die fünf Klassen (ausbalancierte Sampleanzahl mit „random under-sampling“) erreicht werden (siehe Tabelle 7.4).

	1	2	3	4	5	Treffergenauigkeit
1 (gut)	172	32	26	20	18	64,18 %
2	11	174	53	24	6	64,93 %
3	15	61	158	32	2	58,96 %
4	23	21	35	141	48	52,61 %
5 (schlecht)	18	19	14	34	183	68,28 %
Mittelwert						61,79 %

TABELLE 7.4: Konfusionsmatrix der personenabhängigen SVM-Klassifizierung der emotionalen Valenz mittels Audiodaten. 1340 Samples aus Fenstern mit einer Größe von 4,13 Sekunden ohne Überlappung. Evaluiert mit zehnfacher Kreuzvalidierung. Die vorhergesagten Klassen sind die Tabellenspalten.

Es wurde zusätzlich überprüft, inwiefern die Bewertungen der emotionalen Valenz mithilfe der **physiologischen** Daten vorhergesagt werden können. Am besten schnitt hier wiederum die Fusion aus den bereits bei der Klassifizierung der lokalen Klimazone genannten Merkmalen aus Herzrate, IBI und EDA, ab. Die Fenstergröße und Überlappung waren identisch. Die mittlere, personenabhängige Treffergenauigkeit lag bei 79,68 %. Sie wurde mit einer zehnfachen Kreuzvalidierung mit 930 Samples über fünf Klassen (ausbalancierte Sampleanzahl mit „random under-sampling“) ermittelt. Die Konfusionsmatrix ist in Tabelle 7.5 dargestellt.

Generell zeigen diese Experimente, dass die emotionale Valenz, die einen bedeutsamen Faktor des Wohlbefindens darstellt, durch physiologische und Audiodaten der Umgebung abgeschätzt werden konnte. Physiologische Daten zeigten hierbei allerdings bessere Resultate und konnten robuster erfasst werden. Die Audiodaten waren teilweise durch starke Bewegungs- und Windgeräusche nicht verwertbar. Die teilweise zusätzlich zu den Umgebungsgeräuschen aufgezeichneten menschlichen Stimmen beeinflussten in Tests, bei denen Audibereiche mit Stimmen entfernt wurden, die Erkennungsraten allerdings nicht.

Zusammenfassung

Die für dieses Projekt aufgezeichneten Daten zeigten, dass die physiologischen und Umweltdaten mobiler Geräte ausreichend sind, um mittels maschinellem Lernen Rückschlüsse auf die lokale

	1	2	3	4	5	Treffergenauigkeit
1 (gut)	176	0	7	3	0	94,62 %
2	0	170	2	5	9	91,40 %
3	6	9	163	3	5	87,63 %
4	7	21	13	100	45	53,76 %
5 (schlecht)	9	4	13	28	132	70,97 %
Mittelwert						79,68 %

TABELLE 7.5: Konfusionsmatrix der personenabhängigen SVM-Klassifizierung der emotionalen Valenz basierend auf fusionierten Merkmalen von Herzrate, IBI und EDA. 930 Samples aus Fenstern mit einer Größe von 10 Sekunden mit 240 Sekunden Überlappung. Evaluiert mit zehnfacher Kreuzvalidierung. Die vorhergesagten Klassen sind die Tabellenspalten.

Klimazone und das Wohlbefinden zu ziehen. Der direkte Einfluss der lokalen Klimazonen auf die Physiologie konnte ebenfalls nachvollzogen werden. Die eingesetzten Verfahren waren hierbei auf einem Smartphone in Echtzeit, ohne Internet, lauffähig, sodass diese Informationen für Gesundheitsanwendungen zur Verfügung stehen würden, ohne dass die Daten von einem externen Anbieter verarbeitet werden müssten.

7.3.3 Ambiente Anzeige des Status von Fenstern / Türen zur Verbesserung des Sicherheitsempfindens in der Wohnumgebung

Einleitung

Für das Wohlbefinden ist es für den Menschen von größter Bedeutung, sich in seiner Umgebung wohlfühlen. Insbesondere spielt dies in der Wohnung eine wesentliche Rolle. Das subjektive Sicherheitsempfinden kann nachhaltig beeinträchtigt werden, falls beispielsweise schon einmal in die eigene Wohnung eingebrochen wurde. Dies kann wohl kaum vollständig verhindert werden, allerdings ist es möglich, die Wahrscheinlichkeit zu reduzieren, indem die Bewohner alle Fenster und Türen beim Verlassen der Wohnung schließen. Zusätzlich können vergessene offene Fenster, je nach Ausrichtung, bei plötzlich auftretender schlechter Wetterlage (z. B. Gewitter) dazu führen, dass durch Wind und / oder Regen die Inneneinrichtung beschädigt wird. Das Problem des fehlenden Überblicks über Fenster und Türen tritt insbesondere in größeren Wohnungen oder einem Einfamilienhaus auf. Hinzu kommt, dass es bei einer höheren Bewohnerzahl schwieriger ist nachzuvollziehen, ob Mitbewohner Fenster / Türen geöffnet / geschlossen haben. Insbesondere für Senioren kann die Wahrung des Überblicks eine besondere Herausforderung darstellen, da diese evtl. bereits vergesslicher sind oder unter leichten Formen von Demenz leiden.

Zur Lösung dieser Probleme wird in diesem Abschnitt ein System vorgestellt, das die Bewohner über den Zustand (Status) „offen“ / „geschlossen“ von Türen und Fenster informieren sollte. Für die Systemausgabe wurde ambientes Licht, das von RGB-LED-Lampen erzeugt wurde, als ambiente Anzeige verwendet. Die Lampen wurden in der Wohnung, günstig (opportunistisch), platziert, wo diese Information für Bewohner besonders hilfreich war. Die Lampen sollten sich zudem während den dort üblichen Aktivitäten möglichst häufig im Sichtfeld der Bewohner befinden. Beispielsweise war eine Positionierung in der Nähe der Haustür geeignet, wodurch der Status von Fenstern und Türen direkt vor dem Verlassen der Wohnung wahrgenommen werden konnte. Als Sensorik wurden Fenster- / Türkontakte und Bewegungsmelder eingesetzt, die mit einem bereits bestehenden Smarthome-System verbunden waren. Das System wurde mit zwei Bewohnern eines Einfamilienhauses mit Garage evaluiert, da hier ein Überblick über zahlreiche verschiedene Fenster und Türen besonders hilfreich war. Da jede Wohnung einzigartig ist und das System zusammen mit den Bewohnern daran angepasst werden sollte, wurde ein bewohnerzentrierter Entwicklungsprozess angewandt.

Diese Untersuchung wurde in der eigenen Arbeit (Seiderer, Dang & André, 2017) veröffentlicht.

Verwandte Arbeiten

Ambientes Licht Ambiente Anzeigen zur Informationsausgabe mittels Lichtfarbe oder Lichteffekten erfahren in den letzten Jahren zunehmendes Interesse in der Forschung, wie beispielsweise an den Arbeiten von Gustafsson und Gyllenswärd (2005), Hazlewood et al. (2011), Mateevitsi et al. (2014), Matviienko, Rauschenberger, Cobus, Timmermann, Müller et al. (2015) und Müller et al. (2013) erkennbar ist. Hazlewood et al. (2011) untersuchten kleine, kugelförmige Objekte aus Milchglas, die mit verschiedenen Farben beleuchtet waren, um abzubilden, wie sicher sich Studierende mit dem vermittelten Lehrstoff während Universitätsvorlesungen fühlten. In der Arbeit von Occhialini et al. (2011) wurde ambientes Licht verwendet, um das Zeitmanagement der Besprechungen zu verbessern. Hierbei wurden dynamische Lichtmuster an den Wänden eines Besprechungsraumes genutzt. Ein anderes Beispiel, bei dem Zeitmanagement in einer Laborstudie untersucht wurde, stammt von Müller et al. (2013). Ihr „Ambient Timer“ erinnerte Nutzer unaufdringlich an künftige Termine, indem ambientes Licht eingesetzt wurde, das von RGB-LEDs hinter einem Bildschirm stammte. Um das Bewusstsein über den Stromverbrauch zu schärfen, wurde in der Arbeit von Gustafsson und Gyllenswärd (2005) das „Power-aware Cord“ entwickelt und in einem Wizard-of-Oz-Experiment untersucht. Hiermit sollte der Stromverbrauch von elektronischen Geräten durch Lichtmuster auf dem Stromkabel sichtbar gemacht werden.

Keine der verwandten Arbeiten widmete sich der Nutzung von opportunistisch platziertem ambientem Licht, um Bewohner in der Wohnumgebung über den Status der Wohnung zu informieren. Wohnumgebungen, wie Smarthomes unterscheiden sich von öffentlichen Bereichen oder dem Arbeitsplatz, da sie private Bereiche darstellen. Daher kann sich die Akzeptanz der eingesetzten Technologie unterscheiden und war es notwendig, mit dem eigenen System eine Feldstudie durchzuführen.

Eine generelle Studie zur Akzeptanz von ambienten Mitteilungen in der Wohnung wurde von Vastenburger et al. (2009) durchgeführt. Die Autoren fanden heraus, dass die Akzeptanz vom Wichtigkeitsgrad und der Aufdringlichkeit der Systemmitteilungen abhängen. Wichtigere Benachrichtigungen sollten eher aufdringlich, unwichtige unaufdringlich ausgegeben werden. Beim eigenen System sollten keine kritischen Mitteilungen ausgegeben werden, da es sich primär um Erinnerungen bzw. Statusinformationen handelte. Daher wurde eine unaufdringliche Ausgabe mittels ambientem Licht beim eigenen System eingesetzt.

Leitlinien und Kategorisierung von ambientem Licht Für ambiente Informationssysteme (engl. ambient information systems) können diverse Leitlinien und Klassifikationen in der Literatur gefunden werden. Für ambientes Licht existieren allerdings nur wenige Quellen. Neun Leitlinien, die aus Studienergebnissen abgeleitet wurden, sind in der Arbeit von Matviienko, Rauschenberger, Cobus, Timmermann, Müller et al. (2015) aufgelistet. Das eigene System verwendet gemäß deren Leitlinie „GL4“ eine Ampel-Farbkodierung. Diese Farbkodierung wurde für alltägliche Situationen von den Teilnehmern der Studie der Autoren am häufigsten vorgeschlagen. Entsprechend Ampeln und der Leitlinie „GL1“ ist Rot ein negativer und Grün ein positiver Indikator. Dementsprechend wurde diese Farbkodierung auch für das eigene System genutzt.

Matviienko, Rauschenberger, Cobus, Timmermann, Fortmann et al. (2015) schlugen eine Kategorisierung von ambientem Licht vor und wandten es auf 72 Systeme mit ambientem Licht an. Ihre Kategorisierung ist abgeleitet von einigen Leitlinien und Heuristiken für ambiente Informationssysteme, die in der Literatur gefunden werden können. Im Folgenden wird das eigene System entsprechend eingeordnet. Die übersetzten Begriffe, die in (Matviienko, Rauschenberger, Cobus, Timmermann, Fortmann et al., 2015) verwendet wurden, sind hierbei kursiv gedruckt.

Das eigene System zeigte den aktuellen *Status. Mitteilungen*, die Nutzer ablenken können, waren im System nicht enthalten, da keine besonders kritischen Informationen ausgegeben wurden. Daher

wurde nur die Lichtfarbe geändert und die Helligkeit war konstant. Die Information des Systems war durch die *Farbe* (Status der Fenster / Türen) und die *Position* der Leuchtmittel kodiert. Das System wurde im Kontext „Wohnung“ eingesetzt. Die *Informations-Sensitivität* war *privat*, da nur die Bewohner das ambiente Licht sehen konnten. Aus technischer Perspektive wurden beim System *mehrere* LED-Lampen eingesetzt, die *mehrfarbiges* Licht emittierten. Das System war *stationär*.

Konzept opportunistischer ambierter Mitteilungen

Das Konzept für die ambienten Mitteilungen (engl. ambient notifications) bestand darin, diese so unaufdringlich und trotzdem so verständlich wie möglich zu gestalten. Die Bewohner sollten hierdurch nicht gestört werden und die Mitteilungen auf den ersten Blick verstehen. Das System nutzte daher ambientes Licht, das von RGB-LED-Lampen erzeugt wurde, die an besonders geeigneten / opportunistischen Positionen in der Wohnung platziert wurden. Ambientes Licht kombiniert mit einer solchen Platzierung sollte von den Nutzern nur peripher wahrgenommen werden, jedoch zugleich verständlich sein, um die täglichen Nutzeraktivitäten zu unterstützen. Die ambienten Anzeigen gaben mittels Licht die Zustände der Tür- und Fensterkontakte in der Wohnung aus. Bei einer potenziell hohen Anzahl von Tür- und Fensterkontakten mussten die einzelnen Zustandsinformationen bewertet zusammengefasst und mit entsprechender Farbe kodiert angezeigt werden. Es wurde hierfür die von Ampeln bekannte Farbkodierung genutzt, wobei darauf geachtet wurde, dass die Farben gut unterscheidbar waren. Daher wurde gelb mit orange ersetzt, da gelbliches Licht bereits von normalen, warm-farbigem Leuchtmitteln emittiert wird und somit nicht von der restlichen Beleuchtung unterscheidbar gewesen wäre.

Farbe	Bedeutung / Kategorien
Grün	Alle kritischen / unkritischen Türen / Fenster sind geschlossen
Orange	Mindestens eine unkritische Tür oder ein unkritisches Fenster ist offen / gekippt
Rot	Mindestens eine kritische Tür oder ein kritisches Fenster ist offen / gekippt

TABELLE 7.6: Bedeutung / Kategorien der Farbkodierung der ambienten Anzeigen.

Tabelle 7.6 zeigt die Farbkodierung die den aggregierten Zustand der Tür- / Fensterkontakte mit den Kategorien „kritisch“ und „unkritisch“ zusammenfasst. Bei Kontakten, die der Kategorie „kritisch“ zugeordnet wurden, wäre es für Einbrecher leicht, in die Wohnung einzudringen. Hierzu zählten beispielsweise die Wohnungstür oder ein Fenster im Erdgeschoss. Als unkritische Kontakte wurden beispielsweise Fenster im Obergeschoss eingestuft, die zwar geschlossen werden sollten, einen einfachen Einbruch allerdings nicht ermöglichen würden. Solche Fenster werden häufig offengelassen, um beispielsweise ein Schlafzimmer zu lüften oder die Luftfeuchtigkeit im Badezimmer zu reduzieren.

Bewohnerzentrierter Entwicklungsprozess

Die meisten Wohngebäude unterscheiden sich hinsichtlich Architektur, Konstruktion sowie der Anzahl an Stockwerken, Fenstern und Türen. Hinzu kommt, dass jede Wohnung bzw. jeder Haushalt einzigartig sind, da jeder Bewohner eigene Gewohnheiten und tägliche Abläufe besitzt. Wie Hughes et al. (2000) feststellten, ist Technologie in Haushalten normalerweise so konfiguriert und integriert, dass die Abläufe der täglichen Aktivitäten angemessen unterstützt werden. Hierauf zielt auch die bewohnerzentrierte Entwicklung ab. Sie zeichnet sich dadurch aus, dass der Fokus des Designprozesses auf den täglichen Abläufen und damit verbundenen Erfahrungen liegt. Dies steht im Gegensatz zum nutzerzentrierten Ansatz, bei dem Aufgaben und deren Effizienz primär betrachtet werden. Ein ambientes System muss daher so entwickelt werden, dass die entsprechende Wohnung und die täglichen Routinen der Bewohner berücksichtigt werden, um den Bewohnern einen wirklichen Nutzen zu bringen (Vastenburg et al., 2009). Die Entwicklung des hier

vorgestellten Systems folgte daher einem Designprozess, in dessen verschiedene Phasen (Design, Implementierung, Installation) die Bewohner direkt einbezogen wurden.

Bewohner Für die Entwicklung des Systems wurde ein Haushalt akquiriert, der bereits ein Hausautomationssystem mit Bewegungsmeldern in einigen Räumen und Kontaktsensoren an allen Fenstern und Türen besaß. Das Einfamilienhaus bestand aus drei Stockwerken (Keller, Erdgeschoss, Obergeschoss), einem Garten und einer Garage. Der Hausstand setzte sich aus zwei Bewohnern zwischen 60 und 70 Jahren zusammen.

Designprozess Bei einem ersten Treffen wurde das Konzept mit den Bewohnern besprochen und dabei beachtet, dass die Farbkodierung vollständig verstanden wurde. Anschließend wurden Situationen und tägliche Routinen ermittelt, bei denen die Bewohner ein besonderes Interesse zeigten, den Status (offen / geschlossen) der Fenster und Türen zu kennen. Hierzu zählten folgende Situationen / Routinen: • kurz vor dem Verlassen des Hauses, • beim zu Bett gehen und • beim Aufenthalt im Wohnzimmer.

Abhängig von den identifizierten Routinen wurden **Positionen** für die ambienten Anzeigen ausgewählt, um bei den alltäglichen Aufgaben unterstützend wirken zu können. Entsprechend waren Orte, an denen die Bewohner häufig während ihren täglichen Routinen vorbeikamen geeignet. Zu den Orten zählten die Haustür, das Wohnzimmer und Schlafzimmer. Wenn Bewohner sich in der Nähe der Haustür befinden, könnten sie beabsichtigen, das Haus zu verlassen und möchten sicherstellen, dass alle Fenster und Türen geschlossen und abgesperrt sind. Für diese Routine wurde die RGB-LED-Lampe zur Ausgabe von ambientem Licht im Korridor in der Nähe der Haustür installiert (siehe Abbildung 7.12 a)), da sich hier bereits zuvor eine normale, warmweiße LED-Lampe befand. Diese Lampe wurde bereits als automatische Beleuchtung in der Nacht verwendet und beleuchtete teilweise zusätzlich die Treppenstufen in den Keller sowie die in den ersten Stock. Dadurch, dass für die ambiente Anzeige eine Position genutzt wurde, an der bereits ein Leuchtmittel installiert war, stellte sie keine auffällige Veränderung für die Bewohner dar.

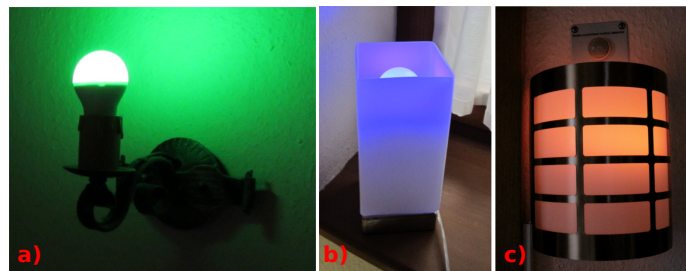


ABBILDUNG 7.12: a) zeigt die RGB-LED-Leuchte in der Nähe der Haustür. Die Lampe in b) war im Wohnzimmer platziert. Lampe c) befand sich in der Nähe des Schlafzimmers. In diesem Fall ist der drahtlose Bewegungsmelder auf der Lampe positioniert.

Wenn die Bewohner sich länger im Wohnzimmer aufhielten (z. B. beim Fernsehen), wollten sie erinnert werden, falls sie vergessen hatten, die Haustür oder ein Fenster im Erdgeschoss zu schließen. Anstatt die Lampen der Deckenbeleuchtung zu ersetzen, wählten die Bewohner eine unaufdringlichere Position in der Nähe der Terrasse aus (siehe Abbildung 7.12 b)). Obwohl sich diese Stelle nicht immer im direkten Sichtfeld der Bewohner befand, konnte die Lampe dennoch leicht wahrgenommen werden. Da eine der täglichen Routinen fernsehen war, musste die Helligkeit der ambienten Anzeige gedimmt werden. Ansonsten hätte sie, insbesondere am Abend, störend gewirkt.

Vor dem Schlafengehen wollten die Bewohner Gewissheit haben, dass keine Tür bzw. kein Fenster mit Sicherheitsrelevanz noch geöffnet waren. Daher wurde die letzte ambiente Anzeige in der Nähe der Treppe im ersten Stockwerk platziert, wo die Bewohner ihr Schlafzimmer hatten

(siehe Abbildung 7.12 c)). Die Helligkeit wurde ebenfalls gedimmt, da es auch als automatisches Nachtlicht diente und die Treppe zum Erdgeschoss beleuchtete.







Der letzte Schritt im Designprozess bestand darin, den Zustand (offen / geschlossen) der Türen und Fenster den Kategorien in Tabelle 7.6 zuzuordnen und somit die **Farbkodierung** festzulegen. Abhängig von der Position (und vom Kontext) einer ambienten Anzeige, mussten die Zustände der Türen und Fenster unterschiedlich berücksichtigt werden, dabei jedoch immer noch verständlich für die Bewohner sein. Die Farbe Grün bedeutete immer, dass alle Türen und Fenster geschlossen waren und in der aktuellen Situation keine Sicherheitsrisiken bestanden.

Fenster im ersten Stockwerk waren generell der „unkritischen“ Kategorie zugeordnet, da sie keinen direkten Zugang in das Gebäude von außen ermöglichten. Ihr Zustand war durch die Farbe Orange indiziert, falls mindestens ein Fenster offen war. Die Bewohner sollten diesen Zustand rückgemeldet bekommen, da hierdurch Beschädigungen, z. B. durch Wetterereignisse wie starkem Wind oder Regen, entstehen hätten können. Es existierte eine Ausnahme im Hinblick auf diese Kategorie. Die Bewohner waren es gewohnt mit einem gekippten Fenster zu schlafen, das nicht dazu führen sollte, dass eine orangefarbene Warnung angezeigt wurde. Ansonsten hätten die Bewohner diese Anzeige vor dem zu Bett gehen ignoriert, wodurch andere offene Fenster der unkritischen Kategorie übersehen hätten werden können. Daher zeigte das ambiente Licht in der Nähe des Schlafzimmers grün, falls nur das Fenster im Schlafzimmer offen war.

Die Farbe Rot indizierte, dass mindestens eine Tür oder ein Fenster der kritischen Kategorie geöffnet war. Dies waren z. B. die Haustür oder Garagentür. Bei dieser Kategorie existierte eine Ausnahme für das ambiente Licht in der Nähe der Haustür. Der Zustand der Haustür und Garagentür wurde nicht verwendet, um die Farbe dieses Lichts anzupassen. Normalerweise wurden vor dem Verlassen des Hauses die Haustür und die gegenüberliegende Garagentür geöffnet. Hierdurch wäre der Status der nicht sichtbaren kritischen Türen / Fenster verdeckt worden. Das Licht war daher nur grün, falls ausschließlich die Garagentür und / oder die Haustür offen waren.

Das ambiente Licht im Wohnzimmer verhielt sich wie das Licht in der Nähe des Schlafzimmers, mit einer Ausnahme, die von den Bewohnern explizit gewünscht wurde. Wenn nur die Garagentür und / oder Haustür offen stand, sollte das Licht die Farbe Blau besitzen. Insbesondere die Garagentür wurde zuvor öfters vergessen und die Indikation sollte daher hervorstechen.

In Abbildung 7.13 ist ein Überblick über die Farbkodierung dargestellt.

Farbe	Position	Bedingungen
	1,2,3	Standardfall; alle Türen / Fenster sind geschlossen
	1,2	mindestens eine Tür / ein Fenster der Kategorie "unkritisch" ist offen
	3	mindestens eine Tür / ein Fenster der Kategorie "unkritisch" ist offen; Ausnahme: Schlafzimmerfenster
	2,3	mindestens eine Tür / ein Fenster der Kategorie "kritisch" ist offen
	1	mindestens eine Tür / ein Fenster der Kategorie "kritisch" ist offen; Ausnahme: Haus- / Garagentür
	2	keine Tür / kein Fenster der Kategorie "kritisch" ist offen und die Garagentür und / oder Haustür ist / sind offen

Position 1 = in Nähe von Haustür; 2 = in Wohnzimmer; 3 = in Nähe von Schlafzimmer

ABBILDUNG 7.13: Übersicht über die eingesetzte Farbkodierung.

Systemaufbau

Das bereits integrierte Hausautomationssystem der Studienprobanden bestand bereits aus drahtlosen Tür- und Fensterkontakten sowie PIR-Bewegungsmeldern des Homematic-Systems. An den Positionen der ambienten Anzeigen befanden sich bereits Bewegungsmelder. Um das System gemäß dem vorgestellten, ambienten Licht-Konzept zu erweitern, wurden drahtlose (2,4 GHz) RGBW-LED-Leuchten des Milight-Systems⁵ eingesetzt. Sie stellten eine preiswerte Alternative zum Philips Hue System dar. Trotz kostengünstigerem Preis funktionierten diese hinreichend zuverlässig und produzierten für bestimmte Farben eine höhere Farbsättigung (besonders bei Grün) als die 2016 verfügbare Generation von Hue-Lampen, deren Farben teils sehr dezent ausfielen. Die Helligkeit der Milight-Lampen war allerdings stärker von der Farbe abhängig als bei den Hue-Lampen. In diesem Fall war die Farbsättigung allerdings relevanter als eine stabile Helligkeit, da sie nur als Ergänzung zur normalen Beleuchtung dienten. Die Homematic-Sensoren und Milight-Lampen benötigten zur Ansteuerung eine Bridge (Funkmodul + (W)LAN-Modul) oder ein entsprechendes Funkmodul am steuernden Rechner. Es wurden Bridges verwendet, da diese unabhängig vom Rechner platziert werden konnten. Dies war hilfreich, da die Reichweite der Funkverbindungen begrenzt ist und der steuernde Rechner somit nicht an der Position platziert sein musste, an dem die Funkverbindungen die beste Sende- bzw. Empfangsleistung erzielten. In Abbildung 7.14 ist eine Übersicht des Systems dargestellt.

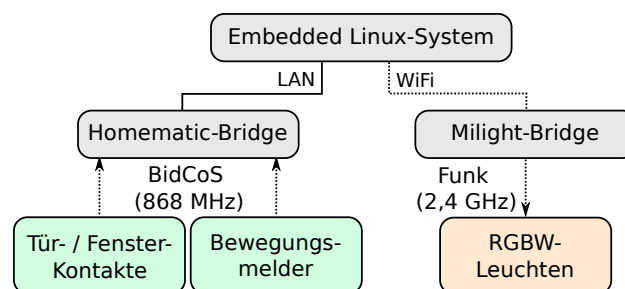


ABBILDUNG 7.14: Übersicht über die relevante Systemhardware.

Als Software zur Ansteuerung der Hausautomations-Hardware wurde die Open-Source-Software FHEM eingesetzt, die auf einem dauerhaft aktiven, energiesparenden Linux-System lief. Das System bestand aus einem Intel NUC-Kit DN2820FYKH mit 4 GiB Arbeitsspeicher und einer SSD mit einer Kapazität von 120 GiB. Die Logik zur Ansteuerung der ambienten Anzeigen wurde in der FHEM-Instanz ergänzt. Sensordaten wurden von FHEM mit Zeitstempeln in Log-Dateien und eine PostgreSQL-Datenbank gesichert.

Einblicke und Diskussion

Nach dem Design, der Implementierung und der Installation des ambienten Lichtsystems wurden Log-Daten gesammelt, die alle Daten der Bewegungsmelder und ambienten Anzeigen beinhalteten. Im Nachfolgenden werden Daten über einem Zeitraum von 6 Monaten ausgewertet. Zusätzlich wurden Interviews mit den Studienteilnehmern durchgeführt.

Das System lief die meiste Zeit stabil und die Funktionalität wurde durch einen Experimentator mindestens zweimal im Monat überprüft. Es traten selten höhere Latenzen zwischen der Erkennung einer Bewegung und der Systemausgabe durch die ambienten Anzeigen auf. Latenzen wurden insbesondere durch die Funkverbindungen der LED-Lampen ausgelöst. In seltenen Fällen wurden die Daten der Kontaktsensoren nicht empfangen, wodurch vorherige Werte für die Ausgabe genutzt wurden. Falls FHEM keine Statusmeldungen von Sensoren über eine längere Zeit erhielt, wurden diese als „tot“ eingestuft und in den Log-Daten gespeichert. Dieses Problem trat beispielsweise mit

⁵<http://www.milight.com/> (16.10.2019)

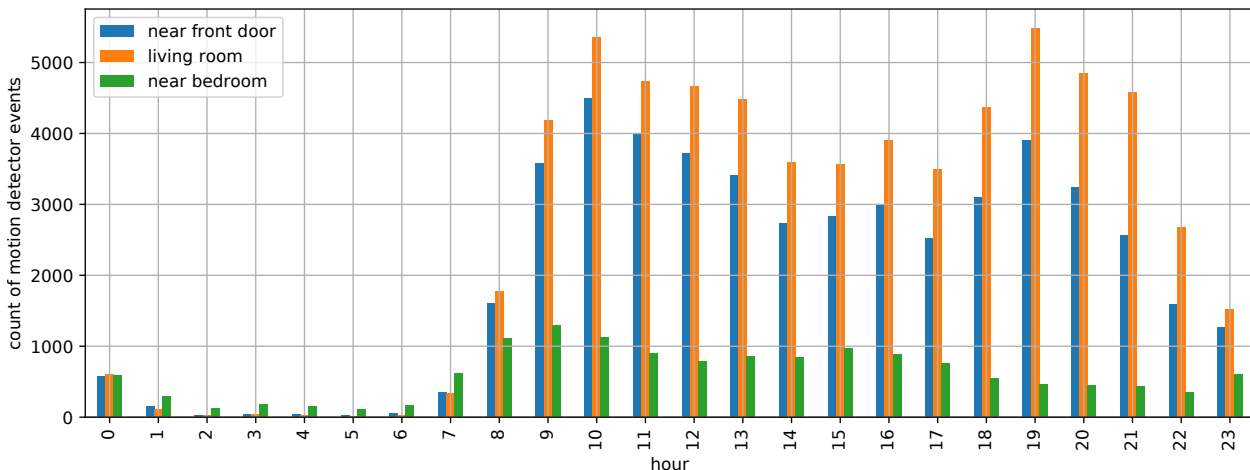


ABBILDUNG 7.15: Summierte Anzahl der erfassten Bewegungen pro Stunde von den Bewegungsmeldern in der Nähe der ambienten Anzeigen. Der dargestellte Zeitraum beträgt 6 Monate.

dem Kontaktsensor an der Garagentür während der Wintermonate auf. Um mögliche fehlerhafte Übertragungen zu den ambienten Anzeigen zu verhindern, wurden die Befehle (z. B. die Farbe des Lichts) bei jeder erkannten Bewegung und bei Änderungen des Zustands einer der Fensterkontakte wiederholt gesendet.

Eine bedeutsame Frage war, wie oft und wann das ambiente Licht für die Bewohner sichtbar war und inwiefern die ausgewählten Positionen geeignet waren und wie häufig sie passiert wurden. Information hierüber konnte durch die Auswertung der Bewegungsmelderdaten gewonnen werden, die in der Nähe der ambienten Anzeigen platziert waren. Erkannte Bewegung führte zur Aktivierung der Lampen. In Abbildung 7.15 ist die Summe der erfassten Bewegungsereignisse der drei Bewegungsmelder pro Stunde über einen Zeitraum von 6 Monaten dargestellt. Zwischen 8 und 0 Uhr wurden die meisten Bewegungen im Wohnzimmer registriert. Die zweithöchste Anzahl konnte in der Nähe der Haustür gemessen werden. In der Nacht und am Morgen zwischen 1 und 7 Uhr passierten die Bewohner hauptsächlich die Lampe in der Nähe des Schlafzimmers. Die ambienten Anzeigen im Wohnzimmer und in der Nähe der Haustür waren für die Bewohner am Tag häufiger sichtbar als die Lampe in der Nähe des Schlafzimmers. Am Abend, wenn die Bewohner schlafen gingen und am Morgen, wenn sie aufstanden, wurde das ambiente Licht in der Nähe des Schlafzimmers wahrscheinlicher von ihnen wahrgenommen.

Die Frage, ob das System einen Einfluss auf das Verhalten der Bewohner hatte, lässt sich nicht zuverlässig mit den aufgezeichneten Daten beantworten. Zwei Voraussetzungen zur Beantwortung der Frage lassen sich allerdings aus den Daten ableiten. Zum einen ist es möglich, abzuschätzen, wann Bewohner das Haus verlassen haben, da nach diesem Zeitpunkt keine Bewegungen im Haus stattfanden. Eine bestimmte Zeitdauer kann definiert werden, ab der die Bewohner als abwesend eingestuft werden, diese wäre allerdings z. B. nachts nicht zuverlässig. Zweitens kann die Farbe der ambienten Anzeigen, die beim Verlassen angezeigt wurde, nachvollzogen werden. Allerdings kann nicht immer festgestellt werden, ob bestimmte unkritische Fenster oder Türen absichtlich nicht geschlossen wurden. Zusätzlich wäre es möglich zu erkennen, ob die Bewohner Fenster oder Türen geschlossen haben und das System somit den Status „grün“ anzeigte, bevor sie das Haus verlassen haben. Allerdings ist es auch hier unklar, ob sie dies aufgrund des ambienten Lichts taten. Zusätzlich sind einige externe Einflüsse möglich, die das normale Verhalten der Bewohner beeinflusst haben könnten und teilweise in den Sensordaten sichtbar sind. Diese Veränderungen können beispielsweise von Besuchern oder schlechtem Wetter über mehrere Tage verursacht worden sein. Auch die Jahreszeiten spielen eine Rolle.

Daten über die mittlere Stundenzahl pro Monat, in denen mindestens ein Fenster oder eine Tür geöffnet waren, sind in Abbildung 7.16 ersichtlich. Die Daten wurden über ca. ein Jahr ermittelt,

wobei das ambiante Lichtsystem erst in den letzten 6 Monaten aktiv war. Es ist deutlich erkennbar, dass in der kalten Jahreszeit die Zeitspannen wesentlich geringer sind, als im Sommer, wo es wärmer und länger hell ist. Die Zeitspanne offener kritischer Türen / Fenster ist generell geringer, besitzt allerdings einen sehr ähnlichen Kurvenverlauf wie bei den restlichen Türen / Fenstern. Ein Rückschluss über die Auswirkungen des Systems ist allerdings nur durch Sensordaten nicht belastbar möglich.

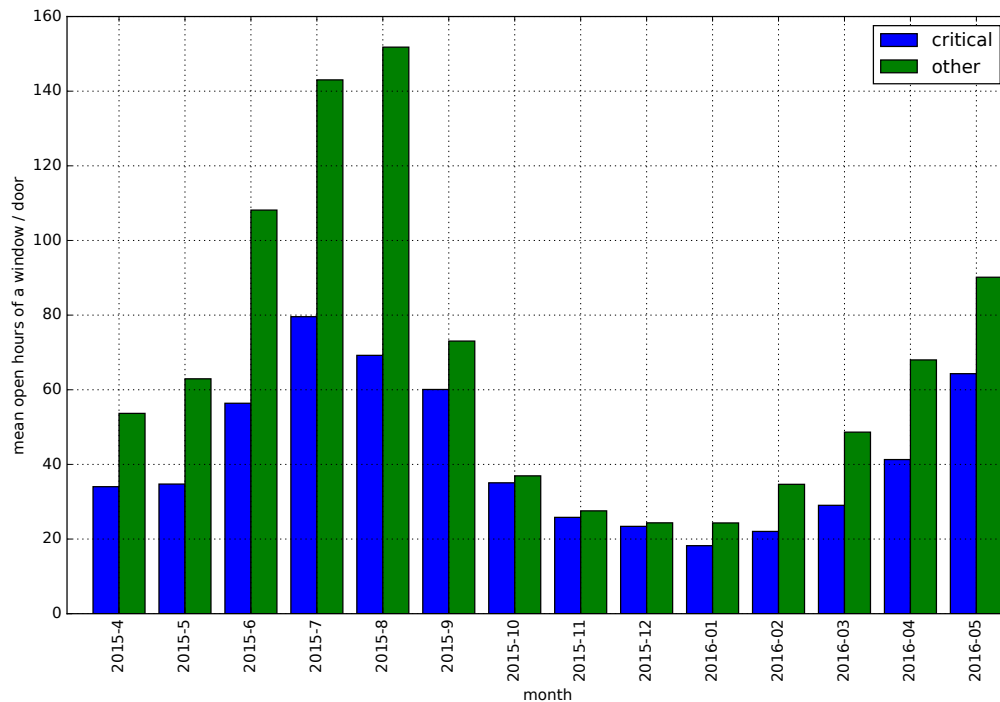


ABBILDUNG 7.16: Mittlere Stundenzahl pro Monat in denen mindestens ein Fenster oder eine Tür offen waren. In der Grafik sind kritische und unkritische Fenster / Türen getrennt aufgeführt.

Daher wurde zusätzlich eine Befragung der Bewohner durchgeführt, inwiefern ihnen das System dabei half, sich an offene Türen und Fenster zu erinnern. Ein Bewohner antwortete sinngemäß darauf, dass sie normalerweise auch selber wüssten, welche Türen / Fenster offen sind, das System ihnen allerdings ein besseres Gefühl gab, auch wenn es teilweise nur ihre Annahme bestätigte. Auch wenn die Lampen nicht direkt anzeigen konnten, welche Tür und welches Fenster offen waren, war es normalerweise ausreichend, die Bewohner daran zu erinnern. In seltenen Fällen mussten sie die entsprechenden geöffneten Türen und Fenster suchen. Die Bewohner bestätigten, dass sie teilweise bewusst ein unkritisches Fenster im Obergeschoss am Tag offen ließen, wenn sie nur für eine kurze Zeit das Haus verließen oder sich im Garten aufhielten.

Zusammenfassung

In diesem Abschnitt wurde ein bewohnerzentrierter Designprozess vorgestellt, der bei der Entwicklung eines Systems angewandt wurde, das mit ambientem Licht unaufdringlich Erinnerungen an offene Fenster / Türen in einem Smarthome mitteilte. Die Resultate der Langzeitstudie wurden in diesem Abschnitt aufgeführt und diskutiert. Die ambienten Anzeigen wurden opportunistisch, unter Berücksichtigung der täglichen Routinen der Bewohner, platziert. Die mit den Bewohnern durchgeführten Interviews legen nahe, dass das System die Anzahl des Auftretens von vergessenen Fenstern und Türen reduzierte. Von größerer Bedeutung war jedoch, dass sich die Bewohner sicherer waren, welchen Öffnungszustand die Fenster und Türen besaßen und einen besseren Überblick über ihre Wohnumgebung erhielten. Daher konnte das entwickelte IoT-System einen positiven Beitrag zu deren Sicherheitsgefühl und somit Wohlbefinden leisten. Da das System mit

zwei Bewohnern getestet wurde, erinnerten sie sich allerdings auch gegenseitig an den Status von Fenstern und Türen im Haus. Daher könnte eine künftige Studie mit alleine lebenden älteren Menschen zeigen, ob diese nicht noch stärker von einem derartigen System profitieren könnten.

7.3.4 Verbesserung der Luftqualität in einem Besprechungsraum

Einleitung

Das Raumklima spielt für das menschliche Wohlbefinden eine erhebliche Rolle. In dieser Untersuchung wurde ein System entwickelt, das in einem Besprechungsraum dabei helfen sollte, eine geringe CO₂-Konzentration beizubehalten. Der Raum verfügte über keine automatische Belüftung, daher mussten die Personen innerhalb des Raumes für ausreichenden Luftaustausch sorgen. Im Raum sollte die Ausgabe des Systems implizit stattfinden, um Besprechungen nicht zu stören. Eine sofortige Reaktion auf schlechte Luft war nicht nötig, da eine höhere CO₂-Konzentration nicht unmittelbar gefährlich ist. Daher eignet sich eine unaufdringliche Systemausgabe in Form von farbkodiertem, ambientem Licht.

Zum Lüften des Raumes ist es notwendig, Fenster zu öffnen. Um zu verhindern, dass danach ein Schließen der Fenster vergessen wird, gab das System eine implizite Erinnerungsanzeige mittels Helligkeitskodierung aus. Hierzu wurden LED-Streifen direkt neben der Tür und den Fenstern platziert, um einen Bezug mittels des Gesetzes der Nähe herzustellen. Das Gesetz der Nähe zählt zu den Gestaltungsgesetzen und besagt, dass Objekte als zusammengehörig empfunden werden, wenn sie einen geringen Abstand voneinander besitzen (Butz & Krüger, 2017, S. 23).

Das entwickelte System sollte relativ kostengünstig (ca. 440 €) sein und problemlos installiert werden können. Die Privatsphäre der Personen im Raum wurde dadurch gewährleistet, indem Sensorik eingesetzt wurde, die keine Rückschlüsse auf bestimmte Personen zulässt und die Daten ausschließlich auf dem Einplatinencomputer innerhalb des Raumes verarbeitet wurden.

Alle Daten der Sensorik wurden über einen längeren Zeitraum aufgezeichnet, um verschiedene Luftqualitätssensoren vergleichen und den Einfluss des Systems auf das Lüftungsverhalten der Raumnutzer auswerten zu können. Es wurde zusätzlich ein Workshop mit fünf Mitarbeitern, die den Besprechungsraum regelmäßig nutzten, durchgeführt.

Die Untersuchung wurde in der eigenen Arbeit (Seiderer, Aslan, Dang & André, 2019) veröffentlicht.

Verwandte Arbeiten

Luftqualität Die Auswirkungen von CO₂ und VOCs auf das menschliche Wohlbefinden wurden bereits in Abschnitt 7.1.1 ausführlich behandelt. Für das entwickelte System wurden die Grenzwerte für CO₂ analog zum Bundesgesundheitsblatt: „Gesundheitliche Bewertung von Kohlendioxid in der Innenraumluft“ (2008) gewählt (< 1000 ppm = gut; 1000-2000 ppm = mittelmäßig; > 2000 ppm = schlecht).

Systeme zur Überwachung der Raumluftqualität In der Literatur können verschiedene Systeme mit ähnlichen Zielen gefunden werden. Einige aktuelle Beispiele werden nachfolgend aufgezählt und mit dem eigenentwickelten System verglichen.

Marques et al. (2019) präsentierten das „iAirCO₂“-System, das Nutzer über „schlechte“ Luft per Mail, SMS oder iOS Smartphone-Meldungen erinnern konnte. Die Daten des im System genutzten NDIR-basierten CO₂-Sensors konnten mittels der Smartphone-App und einem Webinterface dargestellt werden. Der CO₂-Sensor wurde mithilfe eines ESP8266-Mikrocontrollers per WLAN in das System eingebunden. Beim eigenen System wurden ebenfalls ein NDIR-basierter CO₂-Sensor und ESP8266-Mikrocontroller eingesetzt. Die Mikrocontroller dienten hierbei jedoch der Ansteuerung von LED-Streifen, die für die implizite, ambiente Ausgabe benötigt wurden. Für das eigene System spielte insbesondere die unaufdringliche Ausgabe eine wesentliche Rolle, die allen Personen im

Raum ermöglichte, „schlechte“ Luft zu erkennen und selbstständig zu lüften. Auf die Nutzung einer Information für Mitarbeiter des Lehrstuhls über „schlechte“ Luft wie beim „iAirCO₂“-System wurde bei dem eigenen System bewusst verzichtet, da auch externe Personen den Besprechungsraum nutzten, die keinen Zugang zu dieser Information gehabt hätten. Zudem waren Mitarbeiter, die sich nicht im Raum befanden, an entsprechenden Mitteilungen nicht interessiert.

Ein System zur Überwachung der Luftqualität in Schulräumen wurde von Ortiz Perez et al. (2018) beschrieben und in fünf Räumen evaluiert. Die CO₂-Konzentration wurde mit einem neuartigen NDIR-basierten Sensor ermittelt. Zusätzlich wurden die Luftfeuchtigkeit und Temperatur eines Sensorknotens zu einem Gateway übertragen. Es wurde hierbei das Z-Wave-Funkprotokoll eingesetzt. Die überwachten Räume verfügten nicht über eine automatisierte Heizung, Lüftung oder Klimaanlage. Das System wurde zur Sammlung und Analyse von Daten verwendet, um Vorhersagen über die benötigte Lüftungsdauer eines Raumes geben zu können. Die Luftfeuchtigkeit und Temperatur wurden zusätzlich berücksichtigt, da diese das menschliche Wohlbefinden ebenfalls beeinflussen. In ihrer Analyse kamen die Autoren zum Ergebnis, dass es notwendig ist, die Raumluft zu überwachen, damit Schulräume ausreichend gelüftet werden, da dies während ihrer Studie häufig nicht ausreichend erfolgte. In der Studie wurden die Schüler und Lehrer nicht direkt über die aktuelle Raumluftqualität informiert. Die eigene Arbeit könnte das System von Ortiz Perez et al. (2018) diesbezüglich ergänzen und mittels ambienter Anzeigen über den Luftzustand in Echtzeit informieren. Im Vergleich zu Ortiz Perez et al. (2018) wurden beim eigenen System die Temperatur und Luftfeuchtigkeit für das Wohlbefinden nicht berücksichtigt, da diese von Raumnutzern wahrgenommen und teilweise angepasst werden können. Als Erweiterung des eigenen Systems für energieeffizientes Lüften könnten diese Informationen jedoch hilfreich sein.

Ein kommerzielles Produkt zur Überwachung der Raumluftqualität war der „Air Quality Sensor“⁶ von idevices. Das Sensorgerät erfasste die VOCs-, CO₂-Konzentration, Temperatur, Luftfeuchtigkeit und den Luftdruck. Die Daten wurden an den Cloud-Service der Firma übermittelt und gespeichert. Die Sensordaten konnten per Webinterface und Smartphone-App abgerufen werden. Eine Integration in verschiedene Smarthome-Systeme war möglich. Das Gerät verfügte über einen farbigen RGB-LED-Ring zur Anzeige der Luftqualitätsstufen „sehr gut“, „bedenklich“ und „gefährlich“. Hierzu wurde die Ampel-Farbkodierung verwendet, die auch beim eigenen entwickelten System eingesetzt wurde. Das Gerät von idevices besaß ein weißes, rundes Plastikgehäuse ohne Beschriftung oder Symbole. Der RGB-LED-Ring war im Deckel eingelassen und nur von oben sichtbar. Hierdurch war die Bedeutung der RGB-LED-Ausgaben für Raumnutzer nicht ersichtlich. Beim eigenen Prototyp wurde darauf geachtet, dass der Bezug zum Lüften und der Raumluftqualität leichter erkannt werden konnte, indem die ambienten Anzeigen, RGB-LED-Streifen, direkt neben den Fenstern und der Tür platziert wurden. Hierdurch waren die ambienten Anzeigen immer im Sichtfeld von mindestens einer Person. Die Daten wurden vom eigenen Prototypen ausschließlich lokal verarbeitet und gespeichert, wodurch der Schutz der Privatsphäre gewährleistet werden konnte.

Leitlinien und Kategorisierung von ambientem Licht Gemäß den bereits in Abschnitt 7.3.3 eingesetzten Leitlinien von Matviienko, Rauschenberger, Cobus, Timmermann, Müller et al. (2015) wurde eine Ampel-Farbkodierung verwendet (Leitlinie „GL4“). Entsprechend dieser Farbkodierung und der Leitlinie „GL1“ ist Rot ein negativer und Grün ein positiver Indikator.

Mittels der Kategorisierung für ambientes Licht von Matviienko, Rauschenberger, Cobus, Timmermann, Fortmann et al. (2015) wurde auch dieses eigene System eingeordnet. Das System zeigte den aktuellen *Status*. *Mitteilungen*, die Nutzer ablenken können, waren im System nicht implementiert, da die Informationen nie kritischer Natur waren. Daher wurden nur Farbänderungen und ein langsam pulsierendes Helligkeitsmuster verwendet. Die Information des Systems war durch die *Farbe* (Luftqualität), das *Helligkeitsmuster* (Fensterzustand) und die *Position* der LEDs (neben

⁶<https://web.archive.org/web/20191214074546/https://www.idevices.de/> (14.12.2019)

den Fenstergriffen und dem Türgriff) kodiert. Das System wurde im Kontext „Büro“ eingesetzt. Die *Informations-Sensitivität* war öffentlich, da alle Personen im Raum die LEDs sehen konnten. Aus technischer Perspektive wurden beim System *mehrere LEDs* eingesetzt, die *mehrfarbiges* Licht emittierten. Das System war *stationär* im Raum befestigt.

Das hier vorgestellte System unterscheidet sich bezüglich der Kategorisierung zu dem in Abschnitt 7.3.3 durch die zusätzliche Helligkeitskodierung und den Nutzungskontext „Büro“.

Ambientes Licht in der Büroumgebung Ambientes Licht kann in verschiedenen Umgebungen eingesetzt werden. In diesem Fall wurde die Büroumgebung fokussiert, da das eigene System in einem Besprechungsraum an der Universität installiert wurde. Einige Anforderungen überschneiden sich allerdings mit der Heimumgebung. Beispielsweise ist es notwendig, ambientes Licht in Umgebungen mit geringer Helligkeit zu dimmen, damit es nicht störend wirkt. In der Heimumgebung kann dieses Problem am Abend und in der Nacht auftreten, wenn die Bewohner fernsehen und sich ein ambientes Licht in der Nähe befindet. Dieses Problem musste bereits in der eigenen Arbeit aus Abschnitt 7.3.3 behandelt werden und wurde ebenfalls bei ambienten Bildschirmen von Consolvo und Towle (2005) beobachtet. Im Besprechungsraum kann dieses Problem auftreten, wenn Beamer-Präsentationen am späten Nachmittag bzw. Abend gehalten werden oder der Raum hierfür mit Jalousien (innen / außen) abgedunkelt wurde. Daher wurde in das eigene System eine Anpassung der Helligkeit der LEDs an die gemessene Helligkeit im Raum integriert.

Gesundheitsbezogene Systeme mit ambientem Licht, die versuchen, die Büroarbeiter zu mehr Bewegung zu animieren, wurden beispielsweise von Fortmann et al. (2013) und Mateevitsi et al. (2014) vorgestellt. Die Lichtquellen wurden hierbei in der Nähe der Arbeitsplätze positioniert. Biofeedback mit ambientem Licht wurde in der Arbeit von Yu, Hu et al. (2018) untersucht. Hierbei sollte es Stress verhindern und entspannend wirken. In den beiden Arbeiten Müller et al. (2013) und Müller et al. (2015) wurde ambientes Licht hinter den Bildschirmen von Büroarbeitern platziert. In (Müller et al., 2013) wurden die Nutzer dabei an bevorstehende Termine und Treffen erinnert. In (Müller et al., 2015) sollte generell das Verfolgen sekundärer Aufgaben im Hintergrund ermöglicht werden. Die beiden erwähnten Systeme könnten eine Ergänzung zum eigenen System darstellen und ebenfalls das Wohlbefinden bzw. die Zeitplanung der Büroarbeiter verbessern.

Zur Verbesserung des Zeitmanagements in einem Besprechungsraum wurde von Occhialini et al. (2011) ein System mit ambientem Licht untersucht. Weitere Systeme können in der Literatur gefunden werden, die unaufdringlich Informationen während Präsentationen ausgeben können, um den Ablauf von Meetings zu verbessern. Das eigene System sollte jedoch keine direkten Auswirkungen auf die Treffen ausüben und nur unaufdringliches Feedback geben, um den Raum ausreichend zu lüften. Hierdurch sollte ein durch schlechte Luftqualität verringertes Wohlbefinden sowie eine reduzierte kognitive Leistung der Besprechungsteilnehmer verhindert werden.

Implementierung

Hardware Abbildung 7.17 gibt einen Überblick über die verwendete Hardware und deren Kommunikation. Fotos der Hardware zeigt Abbildung 7.18.

Als zentrales Steuergerät wurde ein Raspberry Pi 3 B+ verwendet, der mit dem WLAN verbunden war. Es führte die notwendige Software aus, um Sensordaten zu erfassen, zu verarbeiten und zu speichern. Zusätzlich steuerte es basierend auf Regeln die LED-Aktuatoren. Der Prototyp verfügte über zwei unterschiedliche, kostengünstige CO₂-Sensoren und einen VOCs-Sensor, um deren Messdaten vergleichen zu können. Ferner sollten die Abhängigkeit der Sensoren von Temperatur, Luftfeuchtigkeit und Luftdruck sowie eventuelle Sensordrifts untersucht werden. Für die Ermittlung der Luftqualität war jedoch nur ein zuverlässiger CO₂-Sensor notwendig.

Mit dem Raspberry war ein Arduino Mini Pro verbunden. Er nutzte eine Taktrate von 8 MHz bei einer Spannung von 3,3 V. Der Mikrocontroller wurde verwendet, um einen kalibrierten Sensirion SCD30 NDIR-basierten CO₂-Sensor und Bosch BMP280-Luftdrucksensor per I²C anzusteuern.

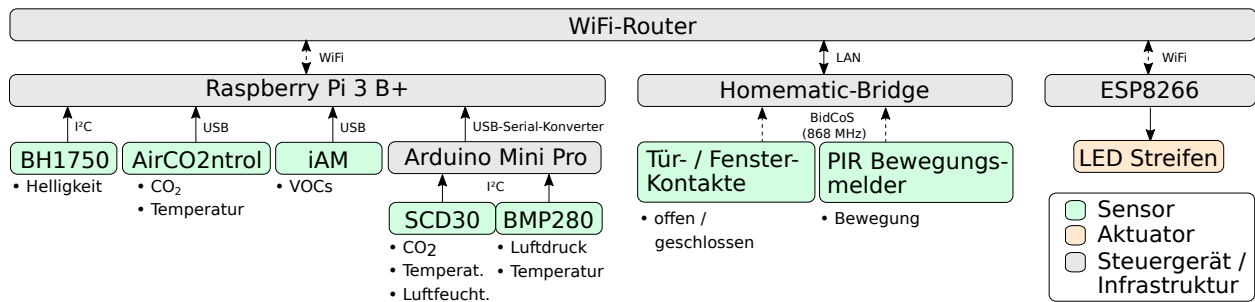


ABBILDUNG 7.17: Überblick über die Hardware des Prototyps.

Der SCD30-Sensor beinhaltet einen kombinierten Temperatur- und Luftfeuchtigkeitssensor, der es der Firmware erlaubt, die Abhängigkeiten bezüglich Temperatur und Luftfeuchtigkeit zu kompensieren. Es war zusätzlich möglich, die Auswirkungen des Luftdrucks auf die Sensordaten auszugleichen, indem der SCD30-Firmware der aktuelle Luftdruck übermittelt wurde. In diesem Fall wurde der Luftdruck vom BMP280-Sensor ausgelesen und vom Arduino an die Firmware des SCD30-Sensors gesendet, sodass dieser an den Arduino bereits vollständig kompensierte Daten liefern konnte. Der Arduino sowie der SCD30- und BMP280-Sensor befanden sich in einem 3D-gedruckten Gehäuse mit Lüftungsschlitzen.

Ein TFA AirCO2ntrol-Mini-CO₂-Sensor, der einen Temperatursensor beinhaltet, war an den Raspberry Pi per USB angeschlossen. Dieses Gerät war kostengünstiger (ca. 40 €) als der neuere SCD30-Sensor (ca. 60 €). Das Gerät wurde so platziert, dass weder das Display noch die LEDs für Raumnutzer sichtbar waren. Laut den Datenblättern konnte der SCD30-Sensor CO₂-Konzentrationen zwischen 400 und 10.000 ppm kalibriert messen, das AirCO2ntrol-Sensorgerät Konzentrationen zwischen 0 und 3000 ppm.

Ein ams iAM VOCs-Sensor war ebenfalls per USB an den Raspberry Pi angeschlossen. Wie viele andere VOCs-Sensoren lieferte der Sensor Werte in ppm, die relativ zur Luftqualität beim Start der Messung oder einem anderen gespeicherten Wert sind. Daher sind bei diesem Sensor nur relative Werte nutzbar.

Um Informationen über den Zustand der Tür und der Fenster sowie Bewegung im Raum zu erhalten, wurden batteriebetriebene Funksensoren des Homematic-Systems eingesetzt. Jedes Fenster und die Tür wurden mit jeweils einem optischen Kontaktsensor (HM-SEC-SCo) ausgerüstet. Ein Bewegungsmelder (HM-SEN-MDIR-SM) befand sich an der Seite des Raums. Die Sensordaten wurden mittels einer Bridge, die mit dem Router per LAN verbunden war, per Funk empfangen.

Die Helligkeit im Raum wurde von einem BH1750-Lichtsensor in Lux bestimmt. Dieser war mit dem Raspberry Pi per I²C verbunden.

Als Aktuatoren wurden sechs (Tür: 1, Fenster: 5) RGB-LED-Streifen verwendet, die jeweils acht WS2812B-LEDs (60 LEDs pro Meter) besaßen und mittels drei ESP8266-Mikrocontroller angesteuert wurden, die mit dem WLAN verbunden waren. Jeder LED-Streifen wurde in ein Aluminiumprofil mit halbdurchsichtiger Plastikabdeckung platziert. Da vergleichsweise hohe Stromstärken benötigt wurden, falls die LEDs weißes Licht mit maximaler Helligkeit emittieren sollten (max. 60 mA pro LED), wurden 5 V-Netzteile mit einer Maximalstromstärke von 2,5 A eingesetzt, um unvorhersehbares Verhalten der Mikrocontroller und LEDs bei instabiler Stromversorgung zu verhindern. Maximal drei LED-Streifen (24 LEDs) waren mit einem ESP8266 und Netzteil verbunden. Die Platinen mit ESP8266, einem Pegelwandler (5 V ↔ 3,3 V) und Anschlüssen wurden von einem 3D-gedruckten Gehäuse geschützt.

Software Abbildung 7.19 zeigt einen Überblick über die Kommunikation zwischen der Software. Auf dem Raspberry Pi war Raspbian-Linux installiert. Zur Kommunikation zwischen einigen Softwarekomponenten wurde der Mosquitto MQTT-Broker verwendet. Zur Datenspeicherung und

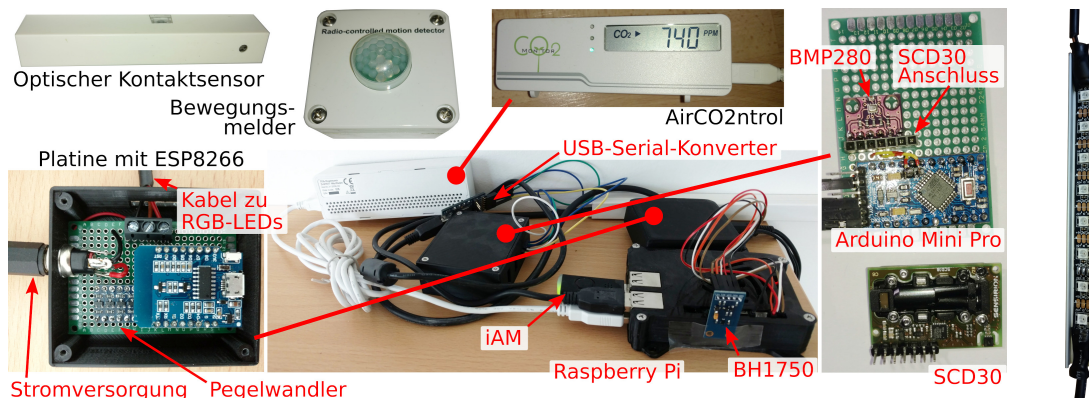


ABBILDUNG 7.18: Das Bild zeigt den Raspberry Pi 3 B+ mit den Gas-, Temperatur-, Luftfeuchtigkeit-, Luftdruck- und Licht-Sensoren. Die Komponenten befanden sich neben der Tür auf einem Sideboard. Zusätzlich ist eine der drei identischen ESP8266-Platinen für die RGB-LED-Streifen, der Homematic-Bewegungsmelder und einer der sechs optischen Kontaktsensoren vom Homematic-System dargestellt. Auf der rechten Seite ist einer der sechs LED-Streifen mit seinen acht LEDs, ohne Abdeckung, abgebildet. Der Streifen ist an ein Aluminiumprofil geklebt.

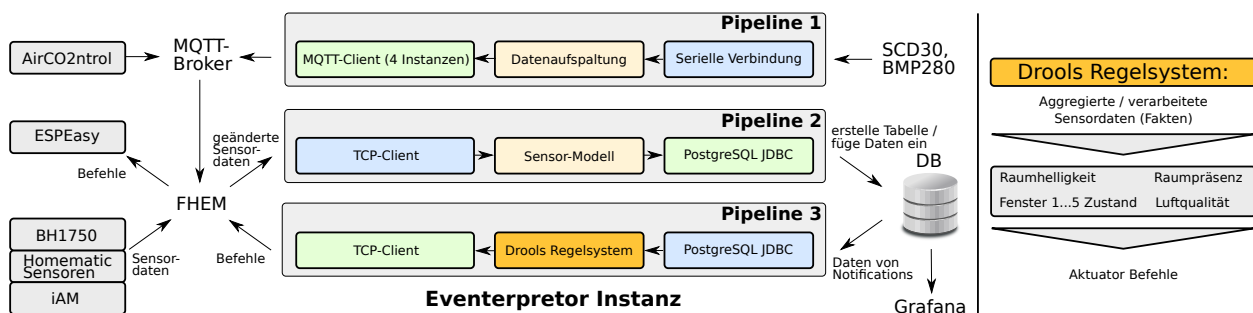


ABBILDUNG 7.19: Überblick über die Datenverarbeitung und Softwarekommunikation.

teilweise auch zur Verarbeitung wurde eine PostgreSQL-Datenbank eingesetzt. Die quelloffene Hausautomations-Software FHEM wurde verwendet, um die Homematic-Sensoren einzubinden und Messwerte des iAM-USB-Sticks und BH1750-Sensors auszulesen. Zusätzlich empfing FHEM alle anderen Sensordaten per MQTT und steuerte die ESP8266-Mikrocontroller an. Eine spezielle Software zur Ansteuerung des AirCO2ntrol-Sensors wurde so angepasst, dass sie die Daten direkt per MQTT versenden konnte. Der Sensor konnte somit in das System direkt eingebunden werden. Der Quelltext ist öffentlich verfügbar⁷. Zur visuellen Exploration und Echtzeitauswertung der Datenbank wurde Grafana eingerichtet.

Um Daten vom Arduino Mini Pro für die Eventverarbeitung und Kontextinterpretation zu erhalten, wurde die eigene quelloffene Software „Eventerpretor“ verwendet. Der Quelltext ist auf Github öffentlich verfügbar⁸. Die Software wurde in Kotlin geschrieben (kompatibel mit Java) und läuft unter Linux und Windows. Sie wurde verwendet, um die Funktionalität von FHEM zu erweitern und die Leistung zu verbessern. Da in der eigenen Software das Sensormodell, die Dateninterpretation und die Regeln implementiert waren, hätte FHEM jederzeit durch eine andere Software ersetzt werden können, die spezielle Hardware wie die des Homematic-Systems direkt ansteuern kann. Hierdurch wäre beispielsweise eine Integration in ein bereits vorhandenes Hausautomationssystem (ohne FHEM) möglich gewesen. Eine Liste mit einigen relevanten Vorteilen bezüglich der Datenverarbeitung gegenüber FHEM sind nachfolgend aufgelistet:

- flow-based Programmierung mit wiederverwendbaren Knoten,

⁷<https://github.com/andreas-seiderer/co2mon>

⁸<https://github.com/andreas-seiderer/Eventerpretor>

- parallele Datenverarbeitung (kann alle vier Kerne des Raspberry Pis ausnutzen),
- datenbankgestützte Datenaggregation,
- unabhängiges Sensormodell und Regelsystem (Drools⁹),
- zwischenspeichern / synchronisieren der Fakten des Regelsystems,
- angepasste Webinterfaces und Wiederholung aufgezeichneter Events für Tests,
- generische Ansteuerung der seriellen Schnittstelle.

Ein Teil der Funktionalität von Eventerpretor ist vergleichbar mit dem komplexeren Node-RED¹⁰ von IBM für Node.js. Allerdings besaß Eventerpretor für den Prototyp benötigte Anpassungen, die in Node-RED nicht verfügbar waren. Zudem war Node-RED, wie auch FHEM, nur beschränkt fähig, Aufgaben parallel zu verarbeiten. Der Zugang zu vielen etablierten Java-Bibliotheken kann ebenfalls als Vorteil von Eventerpretor betrachtet werden. Zwei Webinterfaces (Eingabe und Ausgabe) zur Simulation des vorgestellten Systems sind in Abbildung 7.20 dargestellt. Hierdurch konnten unabhängig von der Hardware die Regeln und das Systemverhalten überprüft werden. Eine automatisierte Prüfung mittels Tests wäre ebenfalls möglich gewesen.

Sensors

Name	Pattern	Color	Brightness
Door	fade	ffd500	100
W1	fade	ffd500	100
W2	constant	ffd500	100
W3	constant	ffd500	100
W4	constant	ffd500	100
W5	constant	ffd500	100

Status

Feed:

```

2019-05-14T13:28:30.799Z    set ESPEasy_ESPRGB_Door_Light raw CANDLE:0:ff0000:100;set
ESPEasy_ESPRGB1_Fenster_1_2 raw CANDLE:0:ff0000:100;set ESPEasy_ESPRGB_Fenster_3_4_5 raw
CANDLE:0:00ff00:100
2019-05-14T13:28:30.794Z    set ESPEasy_ESPRGB_Door_Light raw CANDLE:0:ff0000:100;set
ESPEasy_ESPRGB1_Fenster_1_2 raw CANDLE:0:ff0000:100;set ESPEasy_ESPRGB_Fenster_3_4_5 raw
CANDLE:0:00ff00:100
2019-05-14T13:28:30.795Z    set ESPEasy_ESPRGB_Door_Light raw CANDLE:0:ff0000:100;set
ESPEasy_ESPRGB1_Fenster_1_2 raw CANDLE:0:ff0000:100;set ESPEasy_ESPRGB_Fenster_3_4_5 raw
CANDLE:0:00ff00:100
2019-05-14T13:28:30.796Z    set ESPEasy_ESPRGB_Door_Light raw CANDLE:0:ff0000:100;set
ESPEasy_ESPRGB1_Fenster_1_2 raw CANDLE:0:ff0000:100;set ESPEasy_ESPRGB_Fenster_3_4_5 raw
CANDLE:0:00ff00:100
2019-05-14T13:28:30.798Z    set ESPEasy_ESPRGB_Door_Light raw CANDLE:0:ff0000:100;set
ESPEasy_ESPRGB1_Fenster_1_2 raw CANDLE:0:ff0000:100;set ESPEasy_ESPRGB_Fenster_3_4_5 raw
CANDLE:0:00ff00:100
2019-05-14T13:28:30.799Z    set ESPEasy_ESPRGB_Door_Light raw CANDLE:0:ff0000:100;set
ESPEasy_ESPRGB1_Fenster_1_2 raw CANDLE:0:ff0000:100;set ESPEasy_ESPRGB_Fenster_3_4_5 raw
CANDLE:0:00ff00:100
2019-05-14T13:28:31.731Z    set ESPEasy_ESPRGB1_Fenster_1_2 raw CANDLE:8:00ff00:100
2019-05-14T13:28:31.731Z    set ESPEasy_ESPRGB_Fenster_3_4_5 raw CANDLE:8:00ff00:100
2019-05-14T13:28:31.731Z    set ESPEasy_ESPRGB_Door_Light raw CANDLE:8:00ff00:100
2019-05-14T13:28:34.820Z    set ESPEasy_ESPRGB_Door_Light raw CANDLE:1:00ff00:100
2019-05-14T13:28:34.820Z    set ESPEasy_ESPRGB_Fenster_3_4_5 raw CANDLE:8:00ff00:100
2019-05-14T13:28:34.820Z    set ESPEasy_ESPRGB1_Fenster_1_2 raw CANDLE:8:00ff00:100;set
ESPEasy_ESPRGB1_Fenster_1_2 raw CANDLE:1:00ff00:100
2019-05-14T13:30:03.207Z    set ESPEasy_ESPRGB_Fenster_3_4_5 raw CANDLE:8:ffd500:100
2019-05-14T13:30:03.207Z    set ESPEasy_ESPRGB1_Fenster_1_2 raw CANDLE:8:ffd500:100;set
ESPEasy_ESPRGB1_Fenster_1_2 raw CANDLE:1:ffd500:100
2019-05-14T13:30:03.207Z    set ESPEasy_ESPRGB_Door_Light raw CANDLE:1:ffd500:100
  
```

Sensors

Initialize **Room presence** **Room window1** **Room air quality (CO2 in ppm)**

Init Begin End Opened Closed Set 1210

name presence1 name window1_state name co2_avg

Room window2 **Room brightness**

Opened Closed Set 10

name window2_state name bh1750_avg

Room window3

Opened Closed

name window3_state

Room window4

Opened Closed

name window4_state

Room window5

Opened Closed

name window5_state

Room door

Opened Closed

name door_state

ABBILDUNG 7.20: Webinterfaces von Eventerpretor zur Simulation und Testen des Systemverhaltens. Links sind die Ausgaben des Regelsystems für die LED-Streifen (Farbe, Helligkeit, Helligkeitsmuster) mit den entsprechenden Befehlen dargestellt. Rechts das Interface, mit dem Sensorwerte simuliert werden konnten.

Zur Implementierung der Systemfunktionalität waren beim Prototyp drei Pipelines in einer Instanz von Eventerpretor aktiv. Sie sind in Abbildung 7.19 dargestellt. Die **erste Pipeline** empfangt Daten mittels der seriellen Schnittstelle vom Arduino Mini Pro und verarbeitete diese (z. B. Aufteilung der Werte), sodass sie per MQTT an FHEM gesendet werden konnten.

Die **zweite Pipeline** empfing alle Sensordaten-Events von FHEM. FHEM war so konfiguriert, dass nur veränderte Werte als neues Event weitergeleitet wurden. Die Events von FHEM wurden

⁹<https://www.drools.org/> (18.09.2019)

¹⁰<https://nodered.org/> (18.09.2019)

von Eventerpretor verarbeitet und mit Metadaten angereichert, wie z. B. dem Sensortyp und der Sensorposition. Diese Daten wurden verwendet, um automatisch Tabellen anzulegen, in welche die Sensorwerte eingefügt wurden. Eine Tabelle mit den verknüpften Metadaten aller Sensoren wurde ebenfalls automatisch ergänzt.

Die **dritte Pipeline** empfing verarbeitete Daten direkt von der Datenbank. Dies wurde mittels Triggerfunktionen, die Sensordaten über die Zeit aggregieren, realisiert. Zum Beispiel kann der Mittelwert über die letzten fünf Minuten berechnet werden, sobald ein neuer Sensorwert von der zweiten Pipeline eingefügt wurde. Nach der Berechnung durch die Datenbank wurde eine Notification an alle Subscriber ausgesandt, die den berechneten Wert beinhaltete. Diese Triggerfunktionen wurden von Eventerpretor entsprechend der Konfiguration automatisch erzeugt und es registrierte sich als Subscriber. Andere Programme konnten ebenfalls die Datenbank-Notifications subscriben. Diese Form von asynchroner Benachrichtigung ist eine spezielle Funktion von PostgreSQL und kann sehr hilfreich sein, um ineffizientes Polling zu verhindern. Nach dem Empfang der aggregierten Daten wurden diese in der Pipeline für das Sensormodell und somit für die Fakten des Regelsystems (Drools) verwendet. Das Regelsystem erzeugte zunächst Kontextelemente, wie die Interpretation der Luftqualität aus der CO₂-Konzentration, und erzeugte am Ende die Befehle für die Aktuatoren (siehe Abbildung 7.19 rechts). Die Befehle wurden an FHEM gesendet, das die verschiedenen RGB-LED-Streifen ansteuerte. Ausschließlich neue Befehle wurden versendet, um eine unnötige Belastung des WLANs und der CPU der ESP8266-Mikrocontroller zu verhindern.

Die eigene Firmware des Arduino Mini Pro verwendete zur Ansteuerung des SCD30-CO₂- sowie BMP280-Luftdrucksensors per I²C bereits existierende Bibliotheken. Der ausgelesene Luftdruck wurde zur Kompensation an die Firmware des SCD30-Sensors übermittelt. Der Arduino rief alle Sensordaten des SCD30 (CO₂ in ppm, Temperatur in °C, relative Luftfeuchtigkeit in %) und BMP280-Sensors (Luftdruck in hPa, Temperatur in °C) ab und sandte sie per serieller Schnittstelle an den Raspberry Pi.

Auf die ESP8266-Mikrocontroller wurde die ESPEasy-Firmware¹¹ gespielt. Die Software erlaubte eine problemlose Einrichtung mittels Webinterface, sodass beispielsweise die SSID und das Passwort eines WLANs eingestellt werden konnten, ohne diese fest in den Code einzufügen. Zusätzlich konnten angeschlossene Sensoren und Aktuatoren damit eingerichtet werden. In diesem Fall wurden die WS2812B-LED-Streifen angesteuert, die danach über das Netzwerk gesteuert werden konnten. Es wurde ein vorhandenes Plugin für ESPEasy modifiziert, um neue Helligkeitsmuster nutzen zu können, da diese vom Mikrocontroller gesteuert wurden. Zudem war es nun möglich, mehrere LED-Streifen-Teile an einem ESP8266 unabhängig voneinander anzusteuern.

Aufbau Abbildung 7.21 zeigt die Positionierung der Hardware im Besprechungsraum. In Abbildung 7.22 sind Bilder der LED-Streifen neben der Tür a) und neben vier der fünf Fenster b) und c) dargestellt. Der Raum besaß ein Volumen von ca. 93,4 m³ und wurde von weniger als 20 Personen gleichzeitig verwendet. Durch die Fenster W1, W2 und W3 und nicht zu öffnende Fenster konnte direktes Sonnenlicht in den Raum scheinen. W3 und W4 waren Richtung Norden ausgerichtet.

Falls der Beamer verwendet wurde und zu viel Sonnenlicht in den Raum schien, konnten innere Jalousien und eine äußere, elektrische Jalousie geschlossen werden. Die inneren Jalousien waren manuell mit einer Schnur horizontal bewegbar, der Winkel der Jalousie-Elemente konnte vertikal gesteuert werden. Diese Jalousien befanden sich vor den LED-Streifen, sodass diese nicht mehr sichtbar waren, falls die inneren Jalousien komplett geschlossen wurden. In diesem Fall war nur der LED-Streifen neben der Tür direkt sichtbar.

Der Raspberry Pi, die Gassensoren, der Lichtsensor und der Bewegungsmelder befanden sich auf einem Sideboard, das ca. 110 cm hoch war, sodass die CO₂-Konzentration in einer Höhe gemessen wurde, die ungefähr mit der einer sitzenden Person übereinstimmte. Der Raum hatte fünf Fenster, die geöffnet werden konnten sowie eine Tür, die in einen Flur führte. An der Seite

¹¹<https://www.letscontrolit.com/wiki/index.php?title=ESPEasy> (18.09.2019)

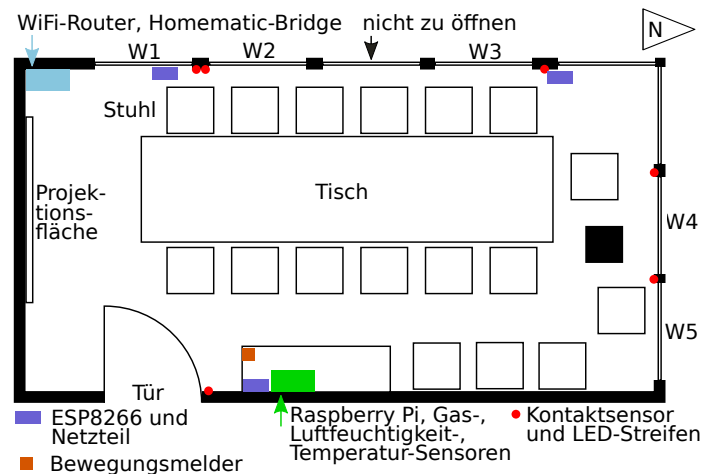


ABBILDUNG 7.21: Überblick über den Besprechungsraum mit den Positionen der Hardware. Zu öffnende Fenster sind mit W1 bis W5 gekennzeichnet.

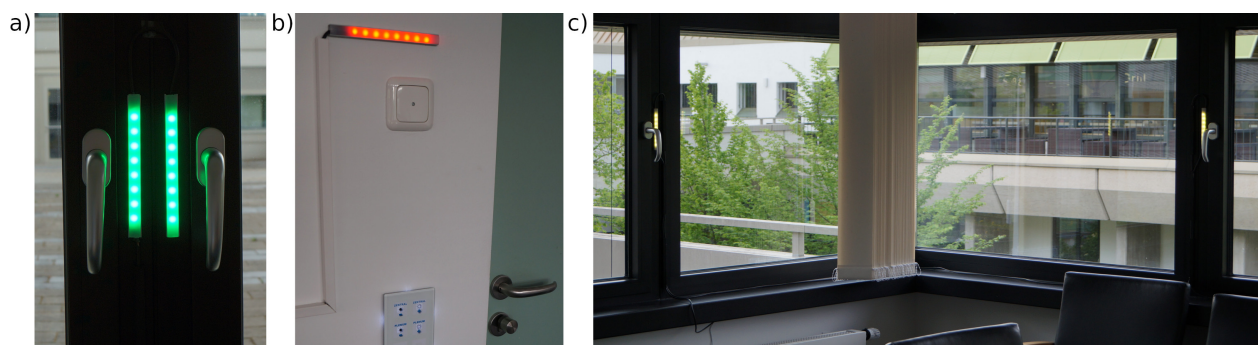


ABBILDUNG 7.22: a) zeigt die LED-Streifen neben Fenster W1 und W2. In b) ist der Streifen neben der Tür dargestellt. In c) sind die Streifen neben Fenster W3 und W4 abgebildet. Zusätzlich sind die beiden nicht zu öffnenden Fenster und eine der manuellen Jalousien (innen) erkennbar. Aus Demonstrationsgründen sind alle drei Farbkodierungen bei den LED-Streifen in den Bildern vorhanden.

jedes Fenstergriffs war ein LED-Streifen platziert. Jede Tür und jedes Fenster wurde mit einem Kontaktsensor versehen.

Evaluation

Für eine grundlegende Evaluation von ambienten Anzeigen und um schwerwiegende Probleme bereits während der Entwicklungsphase zu identifizieren, können die Heuristiken von Mankoff et al. (2003) verwendet werden. Drei instruierte Personen wandten hierzu die Heuristiken auf das System an. Ein Ergebnis hiervon war, dass das pulsierende Helligkeitsmuster verändert wurde. Die niedrigste Helligkeit wurde von null auf einen etwas höheren Helligkeitswert gesetzt, da das Muster zuvor als zu auffällig eingestuft wurde.

Um eine Vorstellung über die zu erwartenden CO₂-Konzentrationen und das Lüftungsverhalten zu erhalten, wurden zunächst Daten ohne Systemausgaben aufgezeichnet. Diese Phase wurde zudem genutzt, um die Systemstabilität zu verbessern. Anschließend wurden die Systemausgaben aktiviert und ein Workshop durchgeführt, um weitere Erkenntnisse über das System zu ermöglichen.

Datenaufzeichnung Erste Datenaufzeichnungen wurden im Winter / Frühling 2018 / 2019 durchgeführt. In Abbildung 7.23 sind exemplarisch Sensordaten dargestellt, die im Folgenden genauer behandelt werden.

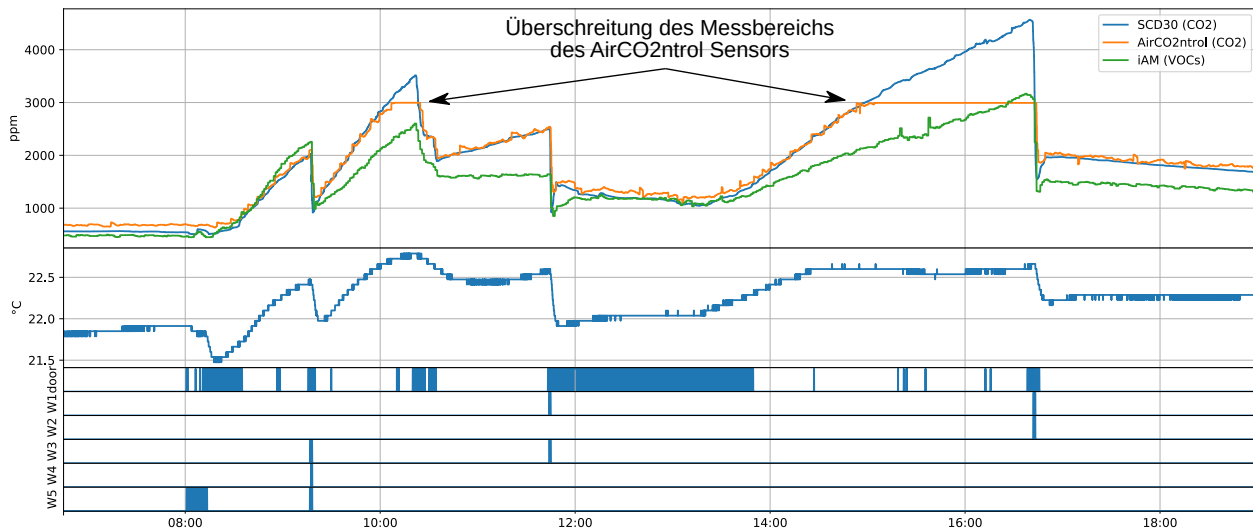


ABBILDUNG 7.23: Beispiel einer Gassensor-Messung im Besprechungsraum an einem Tag im Februar 2019. Die VOCs-Konzentrationen beziehen sich auf frische Luft und können daher nur relativ zu den CO₂-Werten (Werkskalibrierung) betrachtet werden. In diesem Fall sind keine höheren Gaskonzentrationen neben CO₂ ersichtlich, die zu zusätzlichen Ausschlägen in den VOCs-Messungen geführt hätten. Die maximale Außentemperatur lag bei ca. $-1,6^{\circ}\text{C}$. Eine Reduktion der CO₂- und VOCs-Konzentration deutet auf eine offene Tür (door) oder offene(s) Fenster (W1-W5) hin. Der Messbereich des AirCO2ntrol-Sensors wurde zweimal überschritten.

	(1)	(2)	(3)	(4)	(5)	(6)
SCD30 (CO ₂) (1)	1,000	0,993	0,971	0,863	0,972	0,840
AirCO2nt. (CO ₂) (2)	0,993	1,000	0,935	0,934	0,970	0,871
iAM (VOCs) (3)	0,971	0,935	1,000	0,871	0,948	0,839
AirCO2nt. (temp.) (4)	0,863	0,934	0,871	1,000	0,919	0,947
SCD30 (hum.) (5)	0,972	0,970	0,948	0,919	1,000	0,888
SCD30 (temp.) (6)	0,840	0,871	0,839	0,947	0,888	1,000

TABELLE 7.7: Pearson Korrelationskoeffizienten für die Sensordaten, die in Abbildung 7.23 dargestellt sind. Die CO₂-Daten des AirCO2ntrol-Sensors wurden ausgeschlossen, wenn 2900 ppm überschritten wurden (unzuverlässige Messungen).

Die visualisierten Daten des Besprechungsraums wurden an einem Tag im Februar, bei einer maximalen Außentemperatur von $-1,6^{\circ}\text{C}$, aufgezeichnet. Im oberen Bereich der Abbildung sind die Gaskonzentrationen (CO₂ und VOCs), der drei Sensoren ersichtlich. Unterhalb die gemessene Raumtemperatur vom AirCO2ntrol-Sensorgerät. Im unteren Bereich sind die Zustände der Tür (door) und der einzelnen Fenster (W1 - W5) dargestellt.

Die CO₂-Werte des AirCO2ntrol- und SCD30-Sensors unterscheiden sich nur geringfügig. Allerdings ist das Signal des AirCO2ntrol-Geräts stärker verrauscht. Am ausgewählten Tag wurde der Maximalbereich (3000 ppm) des Sensors zweimal überschritten, sodass es generell notwendig ist, Sensoren mit einem höheren Maximalwert, wie den SCD30 zu verwenden, falls die maximalen CO₂-Konzentrationen überwacht werden sollen. Für die Funktionalität des Systems spielten derartig hohe Werte allerdings keine Rolle, da bereits ab 2000 ppm die Farbe auf Rot geschaltet wurde. Die VOCs-Konzentration ist stark von der CO₂-Konzentration abhängig und zeigt hier einen sehr ähnlichen Kurvenverlauf. Allerdings ist dies nicht immer der Fall und würde die Präsenz anderer Gase implizieren. Auch wenn der VOCs-Sensor nur mittels „frischer Luft“ kalibriert wurde, könnte er zusätzlich genutzt werden, um auf potenziell unangenehme Gerüche aufmerksam zu machen.

Tabelle 7.7 zeigt die Pearson Korrelationskoeffizienten, die über die in Abbildung 7.23 dargestellten Daten berechnet wurden. Die CO₂-Daten des AirCO2ntrol-Sensors wurden von den

Berechnungen ausgeschlossen, falls ein Wert von 2900 ppm überschritten wurde, da diese durch ihre geringe Nähe zum Maximum des Messbereichs als unzuverlässig anzusehen sind. CO₂- und VOCs-Konzentrationen zeigen sehr hohe Korrelationen. Auch wenn der SCD30-Sensor gegenüber dem Einfluss von Luftdruck, Temperatur und Luftfeuchtigkeit kompensiert wurde, zeigen sich dennoch, wie bei dem AirCO₂ntrol- und iAM-Sensor, hohe Korrelationen gegenüber diesen Sensordaten. Dies liegt höchstwahrscheinlich daran, dass sich, sobald Personen im Raum waren, die Raumtemperatur und Luftfeuchtigkeit neben den CO₂- und VOCs-Konzentrationen ebenfalls erhöhten. Personen geben zusätzlich zu CO₂ Körperwärme und über ihre Atmung und Haut Feuchtigkeit an die Umgebung ab.

In Abbildung 7.23 ist erkennbar, dass sich die Raumtemperatur an der Messposition während des Tages nur geringfügig (im Bereich von ca. 1,5 °C) verändert hat. Die Gaskonzentrationen und die Raumtemperatur veränderten sich bei geöffneten Fenstern relativ schnell. War ausschließlich die Tür geöffnet, erfolgte eine Reduktion der Gaskonzentrationen wesentlich langsamer.

Generell ist ersichtlich, dass der Raum zwar mehrfach gelüftet wurde, allerdings nicht häufig und teilweise nicht lange genug, wodurch im Tagesverlauf relativ hohe Gaskonzentrationen auftraten. Ein Grund für zu kurzes Lüften könnte der niedrigen Außentemperatur geschuldet sein. Es zeigt sich jedoch insbesondere um ca. 14 Uhr, wo eine längere Raumnutzung stattfand, dass die Personen im Raum, die schlechte Luft scheinbar nicht bemerkten. Der Raum hätte hier gelüftet werden sollen. Die geöffnete Tür sorgte nur sehr begrenzt für niedrigere VOCs- und CO₂-Konzentrationen.

Die Auswertung der Datenaufzeichnung im Winter / Frühling ergab, dass an mehreren Tagen eine schlechte und sogar sehr schlechte Luftqualität im Raum vorhanden war. Jedoch wurden 3000 ppm nur selten überschritten. Die Daten zeigen generell, dass in diesem Raum ein System zur Anzeige der Luftqualität durchaus sinnvoll ist. Die Erfordernis einer Lüftung von Räumen hängt von verschiedenen Faktoren ab, die nur mittels Langzeitmessungen abgeschätzt werden können. Beispielsweise variierte die zeitliche Nutzung des Raumes und dessen Belegung sehr stark. Während der Semesterferien wurde der Raum z. B. deutlich seltener genutzt. Bei überdurchschnittlich kalten oder warmen Außentemperaturen (Winter vs. Sommer) wurde der Raum zudem weniger häufig gelüftet als bei moderaten Temperaturen, da ein Angleichen der Raumtemperatur an die Außentemperatur als unangenehmer empfunden werden kann, als die kaum wahrnehmbare schlechte Luftqualität. Zusätzlich könnte das Lüftungsverhalten auch von den Persönlichkeiten der Nutzer und dem sozialen Gruppenverhalten beeinflusst werden. Möglicherweise überlassen beispielsweise Besucher das Lüften eher den Mitarbeitern oder möchten nicht durch das Öffnen von Fenstern eine Präsentation stören.

Allgemeingültige Aussagen über die Auswirkungen eines Systems auf das Lüftungsverhalten können daher nur sehr eingeschränkt durch die Messung der Luftqualität vor und nach der Installation des Systems ermittelt werden. Es müssen daher auch Raumnutzer befragt werden.

Workshop mit Nutzern des Besprechungsraumes Um Einblicke über das Verhalten und die Meinungen von Nutzern des Besprechungsraumes zu erhalten, wurde mit fünf Mitarbeitern ein Workshop durchgeführt. Die Mitarbeiter verfügten über Expertise in den Bereichen Mensch-Roboter-Interaktion, maschinelles Lernen oder mobile Anwendungen. Drei Forscher führten den Workshop durch. Einer leitete den Workshop, die anderen zwei fertigten Notizen an. Der Workshop wurde im Besprechungsraum durchgeführt, in dem das System installiert war. Zunächst wurde das System den Teilnehmern mündlich erklärt und dessen Funktion anhand einer Demonstration erläutert. Anschließend wurde vom Workshop-Leiter eine semistrukturierte Diskussion mit den Probanden durchgeführt. Diese sollte Informationen über die Einschränkungen, das Potenzial und die generelle Haltung der Teilnehmer bezüglich der Gestaltung der Interaktion liefern. Hierbei standen vier Diskussionsfelder im Vordergrund: 1) Sammeln von Feedback bezüglich des Potenzials und der Schwächen des Systems, 2) zusätzliche Funktionen, die sich Teilnehmer in einem zukünftigen System wünschen würden, 3) Anwendungsfälle, in denen solche ambiente Anzeigen

in anderen Anwendungen und Domänen genutzt werden könnten und 4) eine offene Diskussion über alles, was nicht direkt gefragt wurde. Der Workshop dauerte etwa eine Stunde.

Ergebnisse des Workshops und Diskussion Insgesamt vier Teilnehmer merkten an, dass sie ein derartiges System gerne in ihren Bürozeimern hätten. Nur ein Proband meinte, dass die bereits in ihrem Bürozeimer eine tägliche Routine besäßen, die Fenster zu öffnen, und davon ausgehen, dass hierdurch das Problem gelöst sei. Die meisten der Kommentare der Teilnehmer bezogen sich auf weitere Einsatzmöglichkeiten des Systems. Eine Limitation des Systems wurde ebenfalls angemerkt. Diese würde jedoch nur auftreten, falls die ambienten Anzeigen für verschiedene Arten von Informationen parallel verwendet würden, da in solchen Fällen die Bedeutung der Anzeigen nicht mehr verständlich wäre.

Die wesentlichen Erkenntnisse des Workshops bewegen sich um die Verwendung von Technologie für ambiente Interaktion, die eine neue Form von „ambienter Materialität“ erlaubt. Diese ermöglicht Nutzern eine bessere, materielle Wahrnehmung der Luftqualität, die auch im zukünftigen urbanen Raum eine große Rolle spielt. Normalerweise ist die Luftqualität unsichtbar und nahezu immateriell; sie wird allerdings durch die Nutzung der materiellen LEDs sichtbar gemacht. Im Folgenden werden die Anmerkungen der Teilnehmer zusammengefasst, die sich auf die „ambiente Materialität“ beziehen. Beispielsweise erläuterten Teilnehmer, wie die LED-Streifen genutzt werden könnten, um an verschiedenen Stellen im Raum oder einem Apartment, Empfehlungen zu geben, welche Fenster geöffnet werden sollten und für wie lange, um frische Raumluft zu erhalten. Zusätzlich wünschten sich die Teilnehmer, dass Erinnerungen bezüglich geöffneter Fenster und der Luftqualität an ihre Mobilgeräte gesendet werden. Ein Teilnehmer erwähnte, dass sie daheim im Schlafzimmer normalerweise das Fenster nicht öffnen, da sie vor dem Schlafengehen es häufig vergessen hatten zu schließen und in der Nacht froren. In diesem Fall wäre, laut dem Workshop-Teilnehmer, eine Erinnerung an das Schließen des Fensters, sobald gute Luftqualität erreicht wurde, sehr hilfreich. Allerdings diskutierten die Teilnehmer auch darüber, dass es ein Gleichgewicht zwischen der Luftqualität, Raumtemperatur, Luftfeuchtigkeit und Pollenkonzentration geben muss. Daher wären Sensoren außerhalb der Fenster notwendig, um das Raumklima und die Luftqualität nicht unbeabsichtigt zu verschlechtern.

Ein zweites Thema, das diskutiert wurde, bezog sich darauf, wie digitale Luftqualitätsanzeigen den (urbanen) Raum verändern könnten. Die Workshop-Teilnehmer erwähnten hierzu beispielsweise, dass Informationen über den aktuellen Luftqualitätswert im öffentlichen urbanen Bereich hilfreich sein könnten, um zu entscheiden, welche Route oder welches Personentransportmittel verwendet werden sollte. Ein Teilnehmer konnte jedoch, abgesehen von einer Sensibilisierung der Bevölkerung zum Thema Luftqualität, keinen konkreten Vorteil von öffentlichen Luftqualitätsanzeigen erkennen. Ein weiterer Workshop-Teilnehmer erwiderte hierauf, dass er öffentliche Temperaturanzeigen schätzt, auch wenn diese keinen konkreten Vorteil haben. Die Teilnehmer diskutierten ebenfalls über die Standardisierung von Anzeigen an Fenstern, um die Luftqualität in zukünftigen digitalen Städten zu überwachen und zu verbessern.

Zusammenfassend hat der Workshop gezeigt, dass eine ambiente Luftqualitätsanzeige nicht ausschließlich für einen Besprechungsraum, sondern auch für den urbanen Bereich sehr wünschenswert ist. Die Diskussionen der Teilnehmer haben einige Möglichkeiten, Ideen und Inspiration für zukünftige Forschung aufgezeigt.

Zusammenfassung

In diesem Abschnitt wurden das Design und die Implementierung eines IoT-Systems vorgestellt, das Personen unaufdringlich über die Luftqualität mittels ambientem Licht informiert. Mittels Langzeitaufzeichnungen war ersichtlich, dass die Raumluft in einem Besprechungsraum teilweise hohe CO₂- und VOCs-Konzentrationen aufwies, da nicht oder unzureichend gelüftet wurde.

Hierdurch können u. a. die kognitive Leistungsfähigkeit sowie das generelle Wohlbefinden beeinträchtigt werden. Das entwickelte eigene System konnte somit einen Beitrag zur Verbesserung des Wohlbefindens leisten, da es Raumnutzern ermöglichte, rechtzeitig zu erkennen, wann gelüftet werden sollte. Ein Workshop, der im präparierten Besprechungsraum mit fünf Mitarbeitern durchgeführt wurde, zeigte, dass die Personen die ambienten Erinnerungen verstanden haben und deren Beachtung somit zur Verbesserung der Raumluft führen kann. Beim Workshop wurden zudem einige weiterführende Anwendungszwecke und Anpassungsmöglichkeiten angesprochen. Beispielsweise könnten die Temperatur, Luftfeuchtigkeit, das Wetter und die Luftqualität außerhalb des Raumes zusätzlich berücksichtigt werden, um zu entscheiden, ob eine Raumlüftung empfohlen werden sollte.

7.4 Zusammenfassung des Kapitels

In diesem Kapitel wurden Möglichkeiten zur Erfassung von Umgebungseinflüssen und geeigneter Informationsausgabe untersucht.

Zur Klassifizierung von Umgebungsgeräuschen wurden **tiefe neuronale Netze** eingesetzt, die mittels eines **genetischen Algorithmus** automatisch erstellt wurden. Die erzeugten Netze zeigten gute Klassifikationsergebnisse und konnten mit mehreren normalen Desktop-PCs und GPUs in ca. 5 Tagen erstellt werden. In einer weiteren Untersuchung wurden physiologische und Umweltdaten mittels **mobiler Geräte** mit 7 Personen in 20 Sitzungen auf einem Rundweg von 5 km aufgezeichnet, in der drei lokale Klimazonen lagen (Stadt, Wiese und Wald). Die physiologischen Auswirkungen der Klimazonen konnten mit den Geräten nachvollzogen werden und die Daten des Fitnessarmbands waren ausreichend, um diese mittels maschinellem Lernen klassifizieren zu können. Zusätzlich war es mit diesen Daten möglich, die emotionale Valenz vorherzusagen.

In zwei weiteren Untersuchungen wurden zwei vollständige Systeme für die Wohn- bzw. Büroumgebung entwickelt, die mittels **integrierter Sensorik** und **ambientem Licht** für das Wohlbefinden relevante Umweltdaten erfassten und sichtbar machten. Für eine möglichst ideale Positionierung des ambienten Lichtes sowie für dessen Farb- und Helligkeitskodierung wurden die entsprechenden Wohnungs- bzw. Raumnutzer einbezogen. In der Wohnumgebung stellte das ambiente Licht den Status (offen / geschlossen) von Fenstern und Türen an geeigneten Positionen dar und verbesserte somit das Sicherheitsempfinden der Bewohner. In der Büroumgebung wurde die Luftqualität eines Besprechungsraums für dessen Nutzer mittels ambientem Licht signalisiert. Die ambienten Anzeigen waren hierbei neben den Fenstern platziert, die zum Lüften verwendet wurden. Der Status der Fenster wurde zusätzlich kodiert, um ein Vergessen offener Fenster zu verhindern. In der Regel wird schlechte Luftqualität von Raumnutzern durch Gewöhnungseffekte nicht mehr wahrgenommen. Somit konnte das System dazu beitragen, den Raum ausreichend zu lüften und negative Auswirkungen auf das Wohlbefinden zu reduzieren.

Insgesamt konnten die Untersuchungen in diesem Kapitel zeigen, wie ausgewählte, relevante Umgebungseinflüsse von Systemen zur Verbesserung des Wohlbefindens erfasst und insbesondere in stationären Umgebungen unaufdringlich ausgegeben werden können. Die Nutzung in die Wohnung integrierter Sensorik und ambienter Anzeigen findet sich auch im CARE-Projekt (siehe Kapitel 9) wieder.

8 Physisches Wohlbefinden: Ernährung

In dieser Arbeit wurde aus dem umfassenden Bereich des physischen Wohlbefindens das Thema „Ernährung“ ausgewählt und vertieft. Hierbei wurde die Nahrungsaufnahme einer Person mittels geeigneter Sensorik erfasst und ausgewertet. Weitere Aspekte des physischen Wohlbefindens mit der Zielgruppe Senioren werden in Kapitel 9 behandelt.

Die Ernährung besitzt einen außerordentlich großen Einfluss auf unser Wohlbefinden, unsere Fitness und Gesundheit. Demzufolge ist eine Protokollierung der Nahrungsaufnahme insbesondere dann von Relevanz, wenn Personen unter bestimmten Krankheiten leiden, wie z. B. Diabetes, Übergewicht und Nahrungsmittelallergien. Aber auch für Gesunde ist sie relevant, um eine Unterversorgung mit bestimmten Nährstoffen oder zu hohe Kalorienaufnahme vorzeitig zu erkennen. Das Forschungsgebiet „Ernährung“ gewinnt in der Informatik eine immer größere Bedeutung. Weiterführend kann hierzu beispielsweise der Literaturüberblick von (Altarriba Bertran et al., 2019) herangezogen werden.

In Abbildung 8.1 ist ein vereinfachter Überblick über die in diesem Kapitel enthaltenen Untersuchungen und deren Bezug zum Wohlbefindensmodell dargestellt.

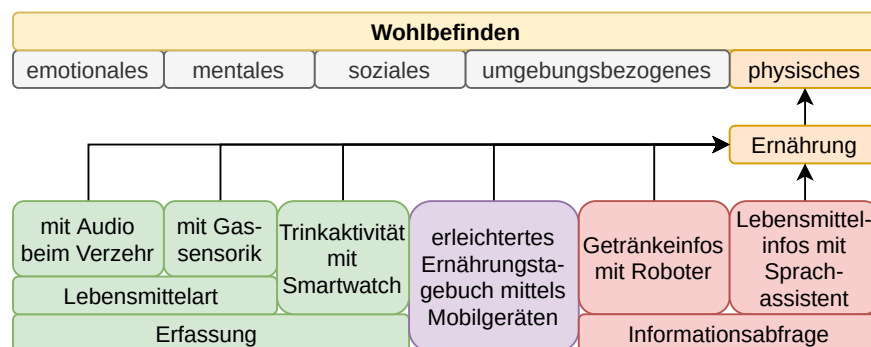


ABBILDUNG 8.1: Überblick über die durchgeführten Untersuchungen in diesem Kapitel.

Erfassung Die Erfassung der Ernährung ist für E-Health-Systeme zur Verbesserung des Wohlbefindens essenziell, um beispielsweise personalisierte Empfehlungen für gesunde Ernährung geben zu können. Folgende wesentliche generelle Probleme wurden in dieser Arbeit bezüglich der Erfassung der Nahrungsaufnahme aus Nutzerperspektive behandelt:

Problem 1: Hoher Aufwand bei der Protokollierung von Nahrungsmitteln

Problem 2: Fehlende Genauigkeit bei der Mengenerfassung von Lebensmitteln

Problem 3: Missachtung der Privatsphäre

Bei einer technischen Umsetzung treten unter anderem beim maschinellen Lernen Probleme auf, die in dieser Arbeit am Beispiel der Lebensmittelerfassung behandelt wurden:

Problem 4: Fehlende Transparenz und Kontrolle beim maschinellen Lernen

Problem 5: Abhängigkeit der Klassifizierungsmodelle von Gassensoren

Das Problem des hohen Aufwandes bei der Nahrungsmittelerfassung (**Problem 1**) wird in den fünf Abschnitten 8.3.1, 8.3.2, 8.3.3, 8.3.4, 8.3.5 behandelt. Um den Aufwand zu reduzieren, werden in dieser Arbeit Verfahren des maschinellen Lernens sowie die Nutzung mehrerer mobiler Geräte (Multi-device interaction) untersucht. Verfahren zur automatischen Klassifizierung der Lebensmittelart wurden mit Audiodaten des Verzehens (siehe Abschnitt 8.3.1) und Gaskonzentrationen (siehe Abschnitt 8.3.2) eingesetzt. Bei der Gassensorik stellen die produktionsbedingten, einzigartigen Eigenschaften der Sensoren (desselben Typs) (siehe Abschnitt 3.5.3) eine besondere Herausforderung bei der Klassifizierung dar (**Problem 5**). Dieses Problem wurde mittels Transferlernen (siehe Kapitel 5) in Abschnitt 8.3.3 behandelt. Als weitere Modalität wurden für die Erkennung von Trinkvorgängen Beschleunigungsdaten von Armbewegungen genutzt (siehe Abschnitt 8.3.5). Da eine automatische Protokollierung der Nahrungsaufnahme nur begrenzt möglich ist, wurde zusätzlich ein System aus mehreren mobilen Geräten entwickelt, mit dem die manuelle Erfassung erleichtert wurde (siehe Abschnitt 8.3.4). Dieses verfügte über einen kompakten Waagen-Prototyp (Smartscale), der das exakte Bestimmen der Lebensmittelmengen erlaubte. Die Smartscale stellte hierbei eine Lösung für eine exakte Erfassung der konsumierten Lebensmittel dar (**Problem 2**). Bei allen in dieser Arbeit eingesetzten Verfahren wurde darauf geachtet, dass die Privatsphäre (**Problem 3**) geschützt wird. Dies wurde bei allen Systemen gewährleistet, da alle Berechnungen auf den Geräten des Nutzers durchgeführt werden konnten und zudem keine Daten unbeabsichtigt die Geräte verließen. Insbesondere bei gesundheitsbezogenen Anwendungen spielt das Vertrauen des Nutzers in ein System eine große Rolle, das durch ungenügende Privatsphäre verspielt werden kann. Maschinelles Lernen ist für den Nutzer häufig intransparent und unkontrollierbar (**Problem 4**). In dieser Arbeit wurde interaktives maschinelles Lernen für die Erfassung von Trinkaktivitäten eingesetzt, bei dem der Nutzer in den Lernvorgang des Systems direkt einbezogen wurde und ein personalisiertes Klassifizierungsmodell erstellt werden konnte (siehe Abschnitt 8.3.5).

Informationsabfrage Die erfassten Daten werden in der Regel von Systemen weiterverarbeitet, um dem Nutzer hilfreiche Informationen und einen besseren Überblick über die eigene Nahrungsaufnahme zu geben. In dieser Arbeit wurden die folgenden generellen Probleme aus der Nutzerperspektive bei der Informationsabfrage behandelt:

Problem 1: Identifizierung ungesunder Lebensmittel

Problem 2: Fehlende Berücksichtigung verschiedener Zielgruppen

Problem 3: Missachtung der Privatsphäre

Für eine gesunde Ernährung ist es von großer Bedeutung, ungesunde Lebensmittel identifizieren zu können (**Problem 1**). In dieser Arbeit werden hierzu zwei Untersuchungen vorgestellt, bei denen Nutzer Informationen zu Lebensmitteln auf einfache Weise abfragen können. Hierzu wurden ein Roboter in Abschnitt 8.3.6 und ein Sprachassistent in Abschnitt 8.3.7 eingesetzt. Der Roboter war Bestandteil eines Systems, das Kalorienangaben zu bestimmten Getränken aus Getränkekartons liefern konnte. Dieses verfügte neben einer Smartscale zur Erfassung der Flüssigkeitsmenge über spezielle Getränkehalter zur Bestimmung des gewählten Getränks. Der Roboter sollte die Informationsabfrage auf einer emotionalen Ebene unterstützen und mit Gesichtsausdrücken seine Sprachausgabe wirkungsvoller gestalten. Die Formulierungen wurden mittels Reinforcement Learning an die Nutzer angepasst und sollten diese dazu bewegen, das gesündere Getränk auszuwählen. Das System wurde während einer Wissenschaftsausstellung präsentiert, wobei die Robustheit gegenüber Fehlern bei der Nutzerinteraktion und Sensorik wesentlich war. Bei dieser Ausstellung mit einer hohen Besucheranzahl war der Lautstärkepegel im Raum hoch, weshalb auf eine Spracheingabe verzichtet wurde.

In einer ruhigen Umgebung, wie der eigenen Wohnung, kann jedoch ein Sprachinterface die Abfrage von Lebensmittelinformationen erleichtern. Daher wurde ein Sprachassistent entwickelt (Abschnitt 8.3.7), der eine zuverlässige Erkennung von verpackten Lebensmitteln per Barcode einsetzte und kurze Fragen zu den Nährwerten mit Bewertungen beantworten konnte. Hierbei wurden zusätzlich besondere Anforderungen für die Zielgruppe Senioren analysiert, da diese besonders von vereinfachten Interaktionsmöglichkeiten bei Computersystemen profitieren können, diese jedoch häufig nicht ausreichend berücksichtigt werden (**Problem 2**). Wie auch bei der Erfassung, wurde bei den beiden Systemen zur Informationsabfrage der Schutz der Privatsphäre berücksichtigt (**Problem 3**). Insbesondere bei den Softwarekomponenten für den Sprachassistenten stellte dies eine Herausforderung hinsichtlich des Gebrauchs der deutschen Sprache dar, konnte jedoch mittels offline nutzbarer offener Software und lokaler Datenbank realisiert werden.

8.1 Nahrungserfassung

8.1.1 Erfassung mittels Mobile Computing und IoT

Die einfachste Form der Lebensmittelerfassung erfolgt mittels Stift und Papier. Diese verursacht allerdings auch den größten Aufwand. Insbesondere die Abschätzung der Menge von Nahrungsmitteln bereitet große Probleme. Für Essen existieren mehrere Hilfsmittel zur manuellen Abschätzung der Portionsgröße (portion size estimation aids: PSEA), die eine höhere Genauigkeit als Fotos ermöglichen, allerdings dennoch hohe Ungenauigkeiten besitzen können, die vom jeweiligen Nutzer abhängen (Sørensen et al., 2014).

Demzufolge wäre eine automatische Erkennung des Nahrungsaufnahmezyklus mit Erfassung des Nahrungsmittels und deren Menge wünschenswert. Hierzu finden sich in der Literatur einige Ansätze, die jedoch jeweils nur Teilbereiche abdecken. Ein kurzer Überblick wird im Folgenden gegeben.

Ein mit Sensorik ausgestattetes Besteck, wie eine Gabel (Kadomura et al., 2013) oder ein Trinkgefäß (Beigl et al., 2001) sind für die Erfassung einsetzbar. Druckveränderungen, die bei der Manipulation von Tellern oder Gefäßen auf einem Tisch entstehen, können mit einer drucksensitiven Tischdecke erkannt werden (Zhou et al., 2015). Um die Armbewegungen zu erfassen, die beim Essen oder Trinken auftreten, wurden von Amft, Junker et al. (2005) Accelerometer verwendet, die auch in einer Smartwatch enthalten sein können (G. M. Weiss et al., 2016). Beim Kauen entstehen Vibrationen, die per Audio erfassbar sind (Amft, Stäger et al., 2005). Für die Erfassung von Essgeräuschen können ebenfalls Mikrofone eines normalen Headsets (Gao et al., 2016) oder einer Smartwatch (Kalantarian & Sarrafzadeh, 2015) genutzt werden. Fotos von Lebensmitteln wurden z. B. in der Arbeit von Ming et al. (2018) zur Erfassung eingesetzt. Spezieller ist die Nutzung von Gassensorik (Hirano et al., 2015; Schroeder et al., 2019).

Während der Vorgang der Nahrungsaufnahme mit verschiedenen Methoden relativ zuverlässig erkannt werden kann, ist es hierbei häufig nicht möglich, die Art des Lebensmittels oder dessen Menge zu bestimmen. Zum Beispiel kann mit Audiodaten ausschließlich, die Konsistenz (z. B. weich / knusprig) des Essens unterschieden werden (siehe Abschnitt 8.3.1 und den Überblick von D. M. Schuller und Schuller (2020)). Allerdings kann nicht erkannt werden, ob die Person beispielsweise sehr fetthaltige oder fettreduzierte Kartoffelchips (mit identischer Konsistenz) zu sich nimmt.

Da auch Menschen visuell viele Lebensmittel und deren Menge unterscheiden können, liegt es nahe, Bilderkennung einzusetzen. Einschränkungen bestehen jedoch auch hier, sodass beispielsweise unstrukturierte Lebensmittel wie Pürees oder Getränke visuell nur begrenzt unterscheidbar sind. Automatisch kann die Bilderkennung nur mit tragbaren Videokameras oder Smartglasses genutzt werden. Dies stellt jedoch außerhalb der eigenen Wohnung aus Gründen der Privatsphäre ein rechtliches Problem dar, wenn nicht zuverlässig verhindert wird, dass Personen aufgezeichnet werden (siehe (Schwenke, 2016) zum Thema Smartglasses). Mit Smartphones besteht dieses

Problem in geringerem Umfang, da hierbei nur gezielt Bilder durch den Nutzer aufgenommen werden. Besonders markante Lebensmittel werden bereits öfters per Smartphone fotografiert und in sozialen Medien geteilt. Zudem existieren Apps zur Lebensmittelprotokollierung bei denen Bilder hinzugefügt werden und neuerdings auch die Lebensmittelart und Zutaten teilweise erkennen können (siehe z. B. Foodvisor¹ oder Snics²). Es handelt sich in diesem Fall jedoch nur um teilautomatisierte Lösungen, da der Nutzer das Bild aktiv erstellen muss. Technisch umsetzbar ist eine Erkennung der Lebensmittelart mittels tiefer neuronaler Netze, wobei eine Bildsegmentierung durchgeführt werden kann, um einzelne Bestandteile wie z. B. eine Zitronenscheibe und ein Schnitzel getrennt zu identifizieren. Transferlernen kann für die Segmentierung und auch Erstellung von Klassifizierungsmodellen besonders hilfreich sein, falls für spezielle Lebensmittel nur wenige Trainings-Daten vorhanden sind (siehe Kapitel 5) oder Modelle personalisiert bzw. für einen Nutzer optimiert werden sollen. Dies kann insbesondere für selbst zubereitete Gerichte notwendig sein. Das Trainieren bzw. „fine-tuning“ entsprechender Modelle mittels Transferlernen ist auch auf den Smartphones direkt möglich. Die allgemeingültigen Untersuchungen aus der eigenen Arbeit (siehe Abschnitt 5.5.2) zum Transferlernen bei der Bilderkennung auf mobilen Geräten können für diesen Anwendungsfall herangezogen werden. Die Bildklassen wären in diesem Fall die Arten von Lebensmitteln bzw. Namen von Gerichten. Diese könnten, ausgehend von einem Basismodell, inkrementell vom Nutzer erweitert und neue Modelle auf dem Smartphone bereits mit kleinen Datenmengen trainiert werden. Modelle zur Erfassung von Lebensmittelarten mittels Transferlernen können erstellt werden, wenn Datenkorpora mit möglichst vielen Nahrungsarten verfügbar sind (z. B. der Datensatz aus (Marin et al., 2019)) oder die Modelle spezialisiert (z. B. Früchte (Mureşan & Oltean, 2018)) werden können. Eine eingeschränkte Bestimmung der Zutaten von Gerichten aus Bilddaten ist mittels tiefer neuronaler Netze ebenfalls möglich (Marin et al., 2019). Falls die Art des Gerichts bereits korrekt erkannt wurde, kann bei der Bestimmung der Menge eine relativ gute Präzision mittels bildbasierter Volumenabschätzung (Jia et al., 2014) erreicht werden.

Falls ausschließlich die Art des Lebensmittels und die Menge erfasst werden sollen, stellt die Erkennung von manuell erstellten Bildern nur eine Alternative zur Auswahl aus Listen, textueller oder Spracheingabe dar. Je nach Leistung der Bilderkennung können diese manuellen Eingabemethoden schneller genutzt werden, da das Erstellen des Fotos, die Erkennung und eventuelle Korrekturen ebenfalls Zeit benötigen. Dies ist insbesondere bei selbst zubereiteten Gerichten der Fall, da für eine hohe Erkennungsrate zunächst eine Personalisierung der Bilderkennung notwendig wäre. Eine Auswahl aus Listen kann beschleunigt werden, indem z. B. eine Kategorisierung vorhanden ist. Bei textueller oder Spracheingabe können entsprechende Lebensmitteldatenbanken zur Autovervollständigung eingesetzt werden und somit den Vorgang stark beschleunigen. Bei verpackten Nahrungsmitteln kann auch der Barcode in Kombination mit einer Datenbank für eine nahezu perfekte Erkennung der Art und Menge herangezogen werden, falls das verpackte Nahrungsmittel nicht verarbeitet und vollständig konsumiert wird.

Eine sehr präzise Bestimmung der Menge eines Getränks oder Gerichts ist durch die Berechnung der Gewichts Differenz zwischen dem befüllten und leeren Gefäß mit einer Waage möglich. Dieser Vorgang kann durch die Nutzung einer Smartscale, die die Gewichtsdaten automatisch übertragen kann, teilautomatisiert werden.

8.1.2 Eigene Arbeiten

Wesentliche Verfahren zur Erleichterung der Erfassung von Lebensmitteln werden in den Abschnitten 8.3.1, 8.3.2, 8.3.3, 8.3.4, 8.3.5 vorgestellt. Eine zusammengefasste Abgrenzung zu anderen Arbeiten wird im Folgenden aufgeführt.

Die im Abschnitt 8.3.1 untersuchte Audioklassifizierung von Essgeräuschen zur Identifikation bestimmter Lebensmittel war Teil der „Paralinguistic Eating Condition (EC) Challenge“ im

¹<https://www.foodvisor.io/en/> (29.09.2020)

²<https://snics.at/> (29.09.2020)

Jahr 2015, bei der unterschiedliche Verfahren von mehreren Gruppen eingesetzt wurden. Die Challenge zeigte, dass es grundsätzlich möglich ist, auch während des Sprechens einer Person zu erkennen, was sie dabei isst. Das in dieser Arbeit präsentierte Verfahren mit einer hierarchischen SVM erreichte, auch im Vergleich zu anderen Methoden, gute Ergebnisse. Es konnten besonders relevante Frequenzbereiche identifiziert und die Erkennungsraten durch die hierarchische Klassifizierung erhöht werden. Eine Nutzung einer hierarchischen Struktur mit aktuellen DNN-Klassifizierungsmodellen könnte auch bei anderen Audioklassifizierungen unterstützend wirken.

Im Abschnitt 8.3.2 wurde kostengünstige, kommerzielle Gassensorik eingesetzt, um verschiedene Lebensmittel zu klassifizieren. Die verwendete Art von Gassensoren stellte für diesen Anwendungszweck eine neuartige Erfassungsmethode dar, da die Sensoren primär zur Erkennung von gefährlichen Gasen entwickelt wurden. Ein wesentlicher Vorteil von Gassensoren besteht darin, dass unverpackte Lebensmittel auch ohne Sichtkontakt erkannt werden können und sich optisch nicht unterscheiden müssen. Die Arbeit baute hierbei auf den Ergebnissen von Hirano et al. (2015) auf. Es wurde in der eigenen Arbeit eine modifizierte Sensorik eingesetzt, die auch einen wesentlich kompakteren und energiesparenden Gassensortyp (MEMS-Fabrikationsverfahren) beinhaltete. Des Weiteren wurden zusätzliche Lebensmittel einbezogen. Die Ergebnisse der Arbeit von Hirano et al. (2015) konnten grundsätzlich nachvollzogen werden. Ergänzend hierzu wurden in der eigenen Arbeit neben Getränken auch Speisen erfolgreich klassifiziert. Zudem konnten hierbei die Klassifizierungen auch mit den Daten von moderneren MEMS-Gassensoren durchgeführt werden. Aktuell existieren auf dem Markt keine neueren, vergleichbaren kommerziellen, kostengünstigen Sensoren. Direkt mit der eigenen Arbeit vergleichbare Experimente / Untersuchungen sind in der Literatur bisher nicht aufzufinden.

Kostengünstige, kommerzielle Gassensorik weist ein einzigartiges Verhalten auf. Somit ist es kaum möglich, Klassifizierungsmodelle mit anderen Sensoren desselben Typs einzusetzen, ohne diese aufwändig mit definierten Gaskonzentrationen zu kalibrieren. Ein Lösungsansatz für dieses Problem wird in Abschnitt 8.3.3 vorgestellt. Hierbei wurde Transferlernen eingesetzt. Das Verfahren basierte auf der Arbeit von Laref et al. (2018). In der eigenen Arbeit wurde das Verfahren allerdings eingesetzt, um ein Klassifizierungsmodell für Lebensmittel mit anderen Sensoren desselben Typs nutzbar zu machen. Für diesen speziellen Anwendungsfall von Transferlernen sind derzeit keine anderen Arbeiten verfügbar. Da für Gassensorik nur kleine Datenmengen zur Verfügung stehen, wäre die Nutzung von tiefen neuronalen Netzen, die bei anderen Problemen gute Ergebnisse zeigen, nur eingeschränkt möglich. Über längere Zeit aufgenommene Datenkorpora können zudem unter dem Problem leiden, dass die genutzte Art von Gassensorik einen altersbedingten Sensor-drift besitzt, der zunächst kompensiert werden müsste, falls dieser zu groß wird. Hierzu wären Langzeituntersuchungen notwendig.

Eine zuverlässige automatische Erfassung der Lebensmittelart und vor allem der Menge ist nicht in allen Situationen möglich. Daher sollte ein entsprechendes System die Möglichkeit besitzen, manuelle Eingaben des Nutzers zu erfassen, die auch zur Korrektur genutzt werden können. Es existiert bereits eine große Anzahl an Apps für Smartphones, die eine manuelle Erfassung erleichtern sollen. Teilweise ist es hierbei auch möglich, Smartwatches zu verwenden. Bei der eigenen Arbeit in Abschnitt 8.3.4 wurde allerdings eine Kombination aus Smartphone, Smartwatch und Smartscale untersucht. Durch die Multi-Geräte-Interaktion und Nutzung verschiedener Modalitäten (Berührung und Sprache) erlaubte es das System, die verschiedenen Vorteile der Geräte, je nach Situation, besser auszunutzen. Verschiedene Erleichterungen bei der Eingabe, wie die Vervollständigung mit einer Lebensmitteldatenbank, wurden umgesetzt. Insbesondere die Ergänzung des eigenen mobilen Smartscale-Prototyps zum Wiegen der Lebensmittelmengen stellte eine Besonderheit dar und erleichterte bei Bedarf die Mengenangabe. Da die Synchronisation (Bluetooth) und Speicherung der Daten vollständig auf den Geräten des Nutzers stattfanden, wurde beim System auf die Wahrung der Privatsphäre geachtet. Das System wurde in einer Feldstudie evaluiert, wobei die Nutzer die Stärken der verschiedenen Geräte in Anspruch nehmen konnten.

In der Literatur sind derzeit für die Nahrungsmittelprotokollierung keine direkt vergleichbaren Gerätekombinationen aufzufinden. Das eigenentwickelte System könnte durch automatische Erfassungsmethoden ergänzt werden. Eine Erweiterung des Systems, um z. B. grafische oder textuelle Auswertungen, wäre ebenfalls möglich. Hierzu gibt es zahlreiche Beispiele in der Literatur und in Form von kommerziellen Anwendungen.

Die in Abschnitt 8.3.5 vorgestellte Software für Smartwatches wurde entwickelt, um Trinkbewegungen automatisch mit dem integrierten Beschleunigungssensor in Echtzeit zu erkennen. Hierfür existierten verschiedene Arbeiten in der Literatur (z. B. G. M. Weiss et al. (2016)). Die Besonderheit bei der eigenen Arbeit liegt darin, dass hierbei interaktives (iML) und kooperatives maschinelles Lernen (cML) genutzt wurden. Das iML fand hierbei in Echtzeit auf einem mobilen Gerät, der Smartwatch, statt. Die Erfassung von Trinkbewegungen stellte hierbei einen Anwendungsfall für die Verfahren dar, die für vergleichbare Probleme einsetzbar wären. Bei iML wird der Nutzer in den Lernvorgang einbezogen, indem dieser dem System Beschriftungen für ausgewählte Daten liefert, die bei der Klassifizierung berücksichtigt werden. In diesem Fall wurden für die Systemeingabe Buttons auf dem Display der Smartwatch eingesetzt. Der Vorteil bei iML besteht darin, dass ein Modell für einen bestimmten Nutzer personalisiert bzw. optimiert werden kann. Zusätzlich erhält der Anwender eine gewisse Kontrolle über den Lernvorgang des Systems. Damit beim iML nicht von Grund auf der Ablauf der Trinkbewegungen erlernt werden muss, wurde ein generisches Basismodell erstellt, wofür ein eigener Datensatz aufgezeichnet wurde. Bei der hierfür notwendigen Datenannotation wurde cML eingesetzt. Mit diesem Verfahren unterstützte die Annotationssoftware den Entwickler, indem es bereits annotierte Bereiche der Trinkaktivität nutzte, um potenzielle weitere Vorkommen im Signal zu markieren, die vom Entwickler überprüft werden konnten. Durch dieses Verfahren konnte die Datenannotation erleichtert werden. Simulationen und eine Nutzerstudie zeigten, dass der Lernprozess beim interaktiven maschinellen Lernen die Erkennungsraten des Systems verbessern konnte. In der Literatur gibt es Beispiele, die iML mit Beschleunigungsdaten einsetzen, um Aktivitäten zu erfassen (z. B. Shahmohammadi et al. (2017)). Bei diesen wurde allerdings das iML nicht direkt auf einer Smartwatch in Echtzeit ausgeführt.

8.2 Abfrage von Informationen zu Lebensmitteln

8.2.1 Interaktionsgeräte

Bei der Abfrage von Informationen zu Lebensmitteln können verschiedene Interaktionsgeräte genutzt werden. Hierzu zählen beispielsweise PCs, Smartphones, Smartwatches / Fitness-Tracker, Tablets und (soziale) Roboter. Im Anwendungsfall der Nahrungsmittelprotokollierung findet die Eingabe meist mittels direkter Manipulation (Begriff siehe Shneiderman (1982)) per grafischem Interface (GUI) mittels Maus und Tastatur oder Berührung statt. Teilweise ist auch eine Spracheingabe möglich. Je nach Gerät können Sensoren eine Eingabe erleichtern. Hierzu zählen beispielsweise eine Kamera zur Erfassung eines Barcodes oder aber auch eine Waage. Die Ausgaben können visuell, auditiv oder taktil erfolgen. Soziale Roboter können, abhängig von ihrem Design, häufig Mimik oder teilweise zusätzlich Gestik mit Kopf und Gliedmaßen ausführen und somit über die Körpersprache Emotionen ausdrücken. Die Fähigkeit sich nonverbal auszudrücken kann die Interaktion mit dem Roboter für Menschen erleichtern (Onyeulo & Gandhi, 2020). In Kombination mit einem Sprachinterface mit natürlicher Sprachverarbeitung ist mit ihnen eine menschenähnliche Interaktion per Konversation möglich. In vielen Fällen wird dies auch von Nutzern erwartet. Sprachassistenten sind unabhängig von der Hardware zu betrachten, da mit ihnen grundsätzlich mittels jeder Hardware, die über ein Mikrofon und einen Lautsprecher verfügt, interagiert werden kann. Beispielsweise kann Amazons Alexa z. B. mit Smartphones, Tablets, Smart TVs, Smart Speakers (Amazon Echo) und speziellen Tablets (Amazon Echo Show) eingesetzt werden. Ihnen fehlt allerdings der visuelle nonverbale Kommunikationskanal, den ein sozialer Roboter besitzt.

Bezüglich der Informationsabfrage zu Lebensmitteln werden alle genannten Interaktionsgeräte, teilweise in Kombination, eingesetzt. Insbesondere für Smartphones ist eine Vielzahl an kommerziellen und offenen Apps verfügbar, die häufig jedoch neben der Abfrage von Lebensmittelinformationen eine Vielzahl an anderen Funktionen besitzen und hierdurch sehr komplex sein können. In der vorliegenden Arbeit stand jedoch eine möglichst einfache Abfrage der Informationen über ein Lebensmittel im Vordergrund. Es wurden hierzu als Interaktionsgeräte ein sozialer Roboter mit eigener, angepasster Sensorik sowie ein Sprachassistent mit einem Tablet eingesetzt.

8.2.2 Eigene Arbeiten

Zwei Systeme, die eine einfache Abfrage von Informationen zu Lebensmitteln erlaubten, werden in den Abschnitten 8.3.6 sowie 8.3.7 behandelt. Hierbei wurden für die entsprechenden Einsatzorte und Zielgruppen unterschiedliche Interaktionsgeräte ausgewählt und angepasste Sensorik genutzt.

In Abschnitt 8.3.6 wurde ein sozialer Roboter bei einer öffentlichen Veranstaltung eingesetzt, der Auskunft über die Kalorienmenge verschiedener Getränke geben konnte. Die Sensorik und der Interaktionsablauf des Systems musste hierbei möglichst robust umgesetzt werden, da in dieser Umgebung keine Spracheingabe (Lautstärke) und Bilderkennung (schlechte Beleuchtung) verwendet werden konnte. Die eigene Arbeit zeigt auf, dass mittels eigens entwickelter Prototypen die Nutzereingaben stabil erfasst werden und für den Reinforcement-Learning-Algorithmus notwendige Daten liefern konnten. Der Roboter kam bei den Besuchern der Veranstaltung gut an, konnte sie allerdings nicht immer zur Wahl von Getränken mit geringerem Kaloriengehalt animieren. Ein autoritäreres Verhalten des Roboters oder ausführlichere Begründungen wären möglicherweise von Vorteil gewesen. In der Literatur gibt es verschiedene Beispiele wie die Arbeit von Tao et al. (2016) bei denen soziale Roboter u. a. für Ernährungsinformationen eingesetzt wurden. Allerdings waren sie nicht mit spezieller Sensorik ausgestattet und Formulierungen wurden nicht an die Nutzer automatisch angepasst.

In Abschnitt 8.3.7 wird das zweite eigenentwickelte System, ein Sprachassistent zur Abfrage von Nährwerten verpackter Lebensmittel, vorgestellt. Das System nutzte hierbei den Barcode und konnte auch Bewertungen zu bestimmten Lebensmittelnährwerten ausgeben. Der Sprachassistent war offline lauffähig und basierte vollständig auf quelloffener Software mit einer offenen Datenbank zum Abrufen der Lebensmittelinformationen. Hierdurch wurde die Privatsphäre geschützt. Das System wurde mit Senioren unter Verwendung eines Tablets als Interaktionsgerät in einem Workshop getestet. In dem Workshop wurden für diese Zielgruppe benötigte Anforderungen ermittelt. Ein Tablet wurde als Hardware für den Sprachassistenten ausgewählt, da diese weiter verbreitet sind als soziale Roboter. Da es sich um ein aufgabenbasiertes Dialogsystem handelte, wurde eine nonverbale visuelle Kommunikation nicht benötigt. Im Vergleich zu Smartphones sind für Senioren zudem Tablets besser geeignet, da sie einen größeren Bildschirm besitzen. Hierdurch können GUI-Elemente größer dargestellt werden, sodass diese leichter mit den Fingern berührt werden können und besser sichtbar sind (siehe hierzu auch Abschnitt 9.1.1). Sprachassistenten zur Abfrage von Informationen lassen sich auch in der Literatur finden (z. B. Shalini et al. (2019)). Diese nutzen allerdings zumeist keine quelloffene Software und insbesondere bei der Spracherkennung, Sprachverarbeitung und Abrufen einer Datenquelle wurden in vielen Fällen Onlinedienste kommerzieller Anbieter eingesetzt, bei denen persönliche Daten über das Internet versendet werden. Beim eigenen Sprachassistenten stellte zudem die Unterstützung der deutschen Sprache eine zusätzliche Hürde bei der Auswahl der quelloffenen Komponenten dar.

8.3 Eigene Untersuchungen

8.3.1 Erkennung von Lebensmitteln beim Sprechen während des Verzehrs mittels Audiodaten

Einleitung

Mittels Audiodaten ist eine grundsätzliche Bestimmung von Lebensmitteln, die während des Sprechens verzehrt werden, möglich. Ein Verfahren der Klassifizierung entsprechender Audiodaten mittels hierarchischen Support Vektor Maschinen (SVM) und Frequenzgewichtung wurde in der eigenen Arbeit (Wagner, [Seiderer](#), Lingenfelder & André, 2015) untersucht. Diese Untersuchung wurde im Rahmen der „Paralinguistic Eating Condition (EC) Challenge“ bei der Interspeech Konferenz 2015 durchgeführt. Ein Audiodatensatz mit sieben Klassen wurde von den Organisatoren bereitgestellt. Er beinhaltet sechs Essgeräusche verschiedener Lebensmittel (Apfel, Haribo „Schlumpfe“, Nektarine, Banane, Chips, Keks), die während des Sprechens einer Person aufgezeichnet wurden und eine Klasse ohne Essgeräusche. Die Aufnahmen stammten von 30 Personen. Details zum Datensatz sind in der Veröffentlichung von Hantke et al. (2016) enthalten.

Ansatz

Abhängig von der Lebensmittelart unterscheiden sich bei deren Verzehr die Audiofrequenzspektren, wobei bestimmte Bereiche von besonderer Relevanz sind und daher hervorgehoben werden können, um eine Klassifizierung zu erleichtern. Bei Untersuchungen von Dacremont (1995) wurde bereits erkannt, dass sich verschiedene Konsistenzen von Lebensmitteln im Audiospektrum des Verzehrvorganges als charakteristisch erweisen. So erzeugen beispielsweise knusprige (engl. *crispy*) Lebensmittel Pegelspitzenwerte im Frequenzbereich oberhalb von 5 kHz und knackige (engl. *crunchy*) Nahrungsmittel (z. B. eine rohe Karotte) im Frequenzbereich von 1,25 bis 2 kHz.

Für alle Tests wurden mit OpenSmile (Eyben et al., 2013; Eyben et al., 2010) die ComParE-Features (6373 Features) (B. Schuller et al., 2015) mit Standardeinstellungen genutzt und SVMs mit linearem Kernel (libSVM: (Chang & Lin, 2011)) für die Klassifikation eingesetzt. Hierbei erfolgten keine Optimierungen der SVM-Parameter. Es wurde die „leave-one-speaker-out“ (LOSO) Kreuzvalidierung verwendet. Alle Experimente wurden mit dem SSI-Framework durchgeführt. Die für die Vergleiche genutzte Baseline besaß eine ungewichtete mittlere Trefferquote (engl. *unweighted average recall = UAR*) von 60,5 %. Die Konfusionsmatrix ist in Tabelle 8.1 abgebildet.

	NO	AP	HA	NE	BA	CR	BI	%
No_Food (NO)	127	1	2	3	5	1	1	90,7
Apple (AP)	5	69	11	29	8	8	10	49,3
Haribo (HA)	2	10	70	13	24	0	0	58,8
Nectarine (NE)	1	33	13	45	31	4	6	33,8
Banana (BA)	5	11	22	27	73	0	2	52,1
Crisp (CR)	1	7	1	4	1	106	20	75,7
Biscuit (BI)	1	18	0	5	3	22	84	63,2
Mittelwert								60,5

TABELLE 8.1: Konfusionsmatrix über die verschiedenen Audioklassen (sechs Arten von Lebensmitteln und kein Lebensmittel). Die vorhergesagten Klassen sind die Tabellenspalten. Die meisten Fehlklassifizierungen traten innerhalb der Teilmengen ({Apple, Haribo, Nectarine, Banana} und {Crisp, Biscuit}) auf. Die entsprechenden Zellen sind in hellgrau markiert.

Hierarchische Klassifizierung Die Erkennungsraten in der letzten Spalte von Tabelle 8.1 zeigen, dass die verschiedenen Klassen unterschiedlich gut erkannt werden konnten. Die Trefferquote von

„Crisp“ (Chips) war beispielsweise wesentlich höher als von „Nectarine“ (Nektarine). Insbesondere traten innerhalb zweier Teilmengen {Apple, Haribo, Banana} und {Crisp, Biscuit} falsche Klassifizierungen auf. Dies ist nachvollziehbar, da die Lebensmittel in den Teilmengen vergleichbare Konsistenz besitzen (eher weich bzw. knusprig). Bei einer Klassifizierung können zunächst diese ähnlichen Klassen zusammen behandelt und in weiteren Klassifizierungen weiter zergliedert werden, bis schlussendlich alle Einzelklassen isoliert wurden. Der Ablauf entsprach einer Baumstruktur. An den Knoten entschieden die verschiedenen Klassifizierer, welcher Ast weiterverfolgt werden sollte, bis ein Blatt, also eine einzelne Klasse, erreicht wurde, die das Ergebnis der Klassifizierung darstellte. Für das Experiment wurde ein Binärbaum verwendet, sodass immer nur zwei Verzweigungen vorhanden waren.

Frequenzgewichtung In Abbildung 8.2 sind die mittleren Frequenzspektren der sechs Klassen, bei denen beim Sprechen Lebensmittel verzehrt wurden, dargestellt.

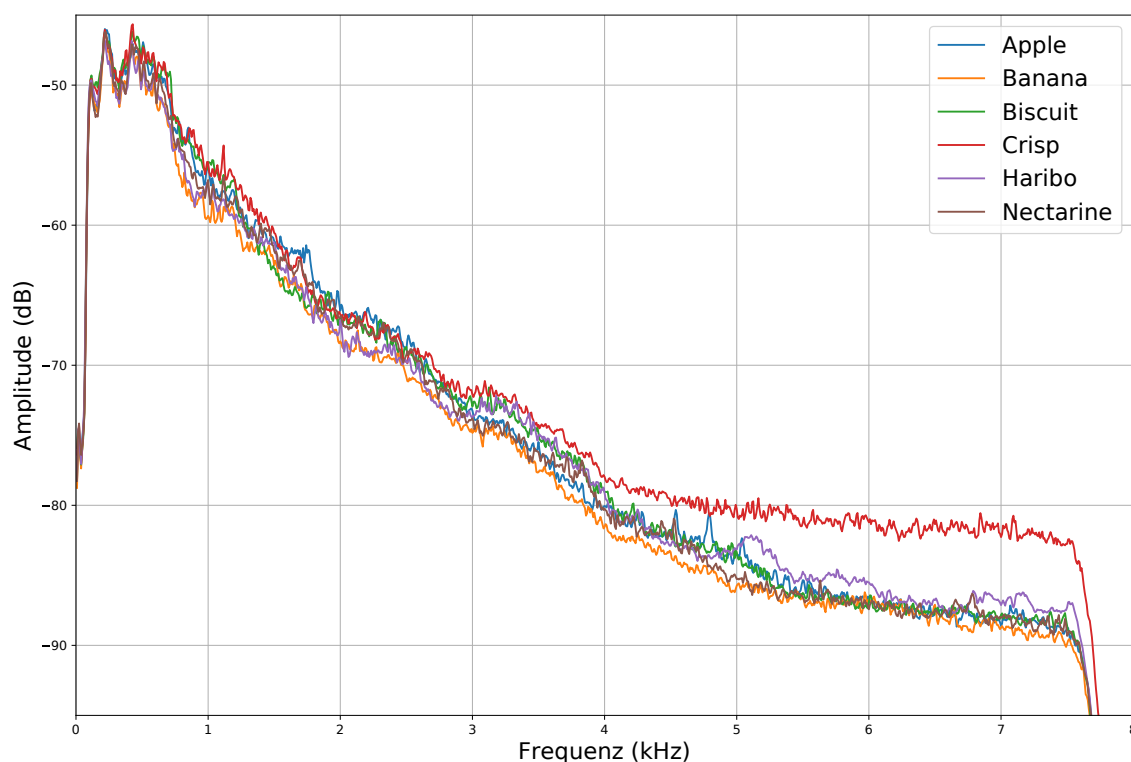


ABBILDUNG 8.2: Mittlere Frequenzspektren der sechs Klassen mit Nahrungsmittel. Die Spektren wurden über alle Dateien im Trainings-Datensatz berechnet.

Im gesamten Frequenzbereich unterscheiden sich die meisten Frequenzverläufe nur geringfügig. Lediglich „Crisp“ sticht heraus, da es ab ca. 4 kHz deutlich höhere Amplitudenwerte besitzt. Dies dürfte von knackenden Geräuschen beim Zerkauen der Chips stammen. Im Gegensatz hierzu sind bei „Banana“ ab ca. 4 kHz insgesamt geringere Amplituden aufzufinden, da das weiche Fruchtfleisch von Bananen zu eher dumpfen, heißt tieffrequenten Geräuschen führt.

Um die Unterschiede zwischen den Klassen zu verstärken, können die dominanten Frequenzbereiche einer Klasse verstärkt und die anderen Bereichen abgesenkt werden. Die Annahme war hierbei, dass dies eine Klassifizierung erleichtern kann. In den Experimenten wurden hierfür zweipolige (zweite Ordnung) parametrische Equalizer (EQ) und zweipolige Butterworth Tiefpass- (engl. lowpass = LP) sowie Hochpass-Filter (engl. highpass = HP) genutzt. Diese Filterimplementierungen stammten aus der Bibliothek SoX³. Die entsprechenden Filterantworten sind in Abbildung 8.3 dargestellt.

³<http://sox.sourceforge.net/> (14.02.2020)

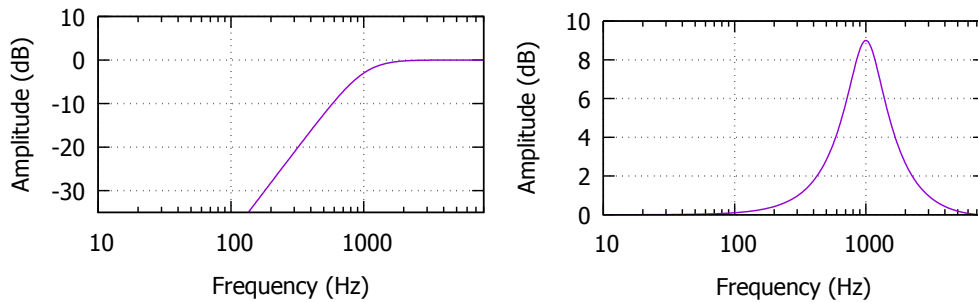


ABBILDUNG 8.3: Links ist der Frequenzgang eines zweipoligen (zweite Ordnung) Butterworth-Hochpass-Filters mit einer Grenzfrequenz von 1 kHz abgebildet. Rechts ist der Frequenzgang eines zweipoligen parametrischen Equalizers (engl. parametric equalizer) mit einer Mittenfrequenz und Bandbreite von 1 kHz sowie einer Amplitudenänderung (engl. gain) von +9 dB dargestellt.

Entfernung stimmhafter Audiodbereiche Beim Anhören und mittels Spektrogrammen konnte erkannt werden, dass Essgeräusche primär in den stimmlosen (engl. unvoiced) Audiodbereichen wahrnehmbar bzw. sichtbar waren. Daher wurden die stimmhaften (engl. voiced) Bereiche entfernt, indem geprüft wurde, ob diese eine harmonische Komponente besitzen. Hierfür wurde eine Implementierung aus OpenSmile eingesetzt.

Ergebnisse

Zunächst wurden die Auswirkungen der einzelnen Ansätze auf die Klassifizierungen untersucht. Anschließend wurden Kombinationen durchgeführt.

Hierarchische Klassifizierung Es wurden drei verschiedene Baumstrukturen getestet (siehe Abbildung 8.4). T-L1 (1 Layer) war die einfachste Struktur, T-L3 die komplexeste. Bei **T-L1** wurde die *NO_FOOD*-Klasse, die als Restklasse am besten erkannt wurde (siehe Tabelle 8.1), durch den ersten Klassifizierer bereits isoliert. Der nächste Klassifikator musste daraufhin ausschließlich die Klassen mit Lebensmitteln klassifizieren. Diese Struktur führte im Vergleich zur eigenen Baseline zu einer leichten Verbesserung der UAR um 0,5 %.

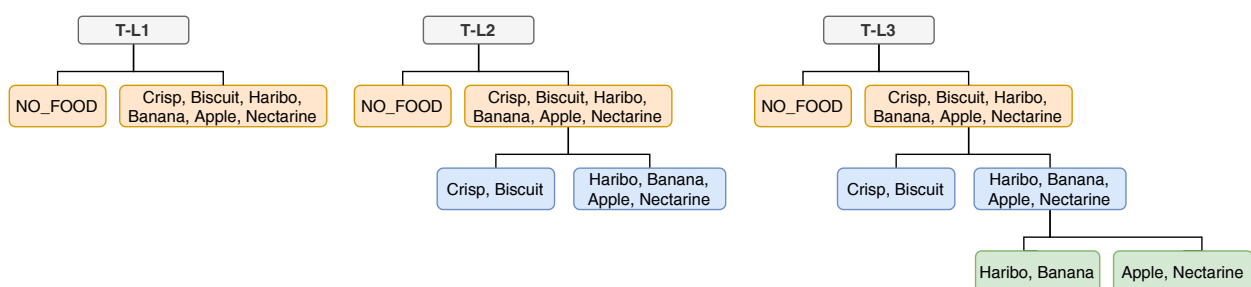


ABBILDUNG 8.4: Die drei getesteten Baumkonfigurationen. Bis auf „NO_FOOD“ wurden die Blätter (Einzelklassen) nicht visualisiert.

Die Klasse mit der zweitbesten Trefferquote war „Crisp“. Diese wurde häufig mit „Biscuit“ verwechselt. Daher wurde bei **T-L2** zusätzlich zu T-L1 eine Klassifizierung zwischen den Gruppen {Crisp, Biscuit} und den restlichen Klassen mit Lebensmitteln durchgeführt, bevor die Einzelklassen mit Klassifikatoren bestimmt wurden. Hierbei trat eine Verbesserung der UAR um 1,2 % auf.

Schließlich wurde die komplexeste Baumstruktur (**T-L3**) umgesetzt, bei der im Vergleich zu T-L2 ein weiterer Klassifikator zwischen {Haribo, Banana} und {Apple, Nectarine} eingeführt wurde, da bei diesen Klassen zwischen „Apple“ und „Nectarine“ die meisten Verwechslungen bei der Klassifizierung auftraten. Hierbei konnte eine Verbesserung der UAR um 2,2 % festgestellt werden.

Frequenzgewichtung Als nächster Schritt wurde die Frequenzgewichtung getestet. Da es schwierig war, bestimmte Frequenzbereiche für alle Lebensmittel zu isolieren, wurden der parametrische Equalizer, Tiefpass- und Hochpass-Filter mit den Einstellungen wie in Abbildung 8.3 in 500 Hz-Schritten über den gesamten Frequenzbereich einzeln angewandt. Der Equalizer und Hochpass-Filter konnten insbesondere zwischen 4,0 und 6,5 kHz gute Ergebnisse erzielen und die UAR um 4,7 % verbessern.

Entfernung stimmhafter Audibereiche In den Experimenten konnte festgestellt werden, dass Essgeräusche tatsächlich in den stimmlosen Audibereichen am besten erkannt werden können. Allerdings konnte nur durch die Kombination der Features der stimmlosen und stimmhaften Bereiche eine deutliche Verbesserung der UAR (um 3,7 %) beobachtet werden.

Kombination der Ansätze Schließlich wurden Kombinationen der verschiedenen Ansätze evaluiert, indem die zuvor besten Ergebnisse genutzt wurden. Es wurde die T-L3 Struktur angewandt sowie die entsprechenden besten Frequenzgewichtungen und Filter ermittelt und eingesetzt (siehe Abbildung 8.5). Insgesamt konnte hiermit eine Verbesserung der UAR um 7,6 % erreicht werden, was das beste Ergebnis der Untersuchungen darstellte. Es wurden zusätzlich die stimmhaften Audibereiche entfernt, allerdings schien hierbei die Frequenzgewichtung keine positiven Auswirkungen mehr zu besitzen. Dies könnte daran gelegen haben, dass im Audiospektrum der stimmlosen Bereiche die Lebensmittelgeräusche bereits ausreichend hervortraten.

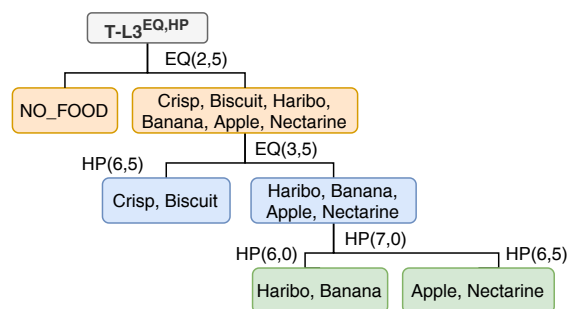


ABBILDUNG 8.5: Die beste, getestete Kombination aus hierarchischer Klassifikation und Frequenzgewichtung. EQ = Equalizer, HP = Hochpass-Filter. Die Zahl in Klammern gibt jeweils die Mittenfrequenz bzw. Grenzfrequenz des Filters in Kilohertz an. Beispielsweise wurde ein Hochpass-Filter vor der Klassifizierung der übrigen Klassen „Apple“ und „Nectarine“ mit einer Grenzfrequenz von 6,5 kHz angewandt.

Bestes Ergebnis der Experimente

Bei den durchgeführten Experimenten zeigte die Kombination aus Baumstruktur und Frequenzgewichtung ($T-L3^{EQ,HP}$) die besten Ergebnisse: 68,1 % UAR (7,6 % Verbesserung) im Vergleich zur eigenen Baseline (60,5 % UAR), berechnet auf dem Trainings-Datensatz. Die offizielle Baseline der Challenge lag hier bei 61,3 % (6,8 % Verbesserung). Bei Nutzung des vollständigen zur Verfügung gestellten Trainings-Datensatzes (Train / Test) lag die UAR von $T-L3^{EQ,HP}$ bei 67,6 %, die eigene Baseline bei 63,5 % (UAR) und die offizielle Baseline bei 65,9 % (UAR).

Insgesamt konnte daher mit dem besten, getesteten Verfahren eine Verbesserung von 6,8 % (Train) bzw. 1,7 % (Train / Test) UAR gegenüber der offiziellen Baseline erreicht werden.

Die einzelnen Ergebnisse der Experimente sind im Papier (Wagner, [Seiderer](#), Lingenfelder und André (2015)) dargestellt.

Zusammenfassung

Die Experimente zeigten, dass es durch die Nutzung einer hierarchischen Klassifizierung, angepasster Frequenzgewichtung und das Entfernen von stimmhaften Audiodbereichen möglich ist, die Klassifizierungsleistung, im Vergleich zu einer einfachen SVM-Klassifikation, zu verbessern. Insgesamt schnitt hierbei die Kombination aus einer hierarchischen Baumstruktur und Frequenzgewichtung am besten ab. Diese Technik scheint daher eine vielversprechende Lösung bei vergleichbaren Problemen darzustellen. Es sollte berücksichtigt werden, dass bei diesen Tests nicht alle Parameter optimiert (z. B. bei der SVM) oder andere Klassifikationsverfahren (z. B. tiefe neuronale Netze) getestet wurden, sodass noch weiteres Verbesserungspotenzial möglich sein könnte.

Generell zeigten die Experimente bei dieser Challenge, dass es möglich ist, sechs verschiedene Lebensmittel beim Verzehren während des Sprechens anhand der Audiodaten zu unterscheiden. Dies ist insbesondere interessant, da diese Unterschiede selbst für Menschen kaum wahrnehmbar sind. Eine praktische Anwendungsmöglichkeit im IoT-Bereich besteht beispielsweise bei einem Sprachassistenten, der während der Interaktion implizit Informationen über die Ernährungsgewohnheiten des Nutzers ermitteln könnte. Beispielsweise könnte er darauf hinweisen, dass es unhöflich ist, während des Sprechens zu essen. Zudem wird durch diese Verhaltensweise die Spracherkennung des Assistenten verschlechtert.

8.3.2 Erkennung von Lebensmitteln mittels Gassensorik

Einleitung

In der eigenen Arbeit (Dang, [Seiderer](#) & André, 2018) wurde untersucht, inwiefern sich kostengünstige Gassensorik dazu eignet, verschiedene Lebensmittel zu klassifizieren. In der Arbeit wurden einige Resultate bezüglich der Erkennung von Getränken des Artikels von Hirano et al. (2015) nachvollzogen und es konnten vergleichbare Trefferquoten und Genauigkeiten erreicht werden. Des Weiteren konnte in der eigenen Arbeit gezeigt werden, dass zusätzlich zu den Getränken verschiedene feste Lebensmittel klassifiziert werden können. Ein wesentliches Ergebnis der eigenen Untersuchungen war zudem, dass die kompakteren und stromsparenderen MiCS-Gassensoren eine gute Alternative zu den älteren MQ-Gassensoren darstellen. Die MiCS-Gassensoren wurden in den Untersuchungen von Hirano et al. (2015) nicht betrachtet.

Für die Experimente wurden in der eigenen Arbeit die zwei stationären Gas-Sensorboxen eingesetzt, die in Abschnitt 3.5.5) beschrieben wurden. Neben den MQ-Sensoren standen daher auch der modernere MiCS-Sensor sowie insgesamt ein VOCs- und CO₂-Sensor zur Verfügung. Die in den Boxen integrierten Licht- und Feinstaubsensoren wurden nicht verwendet. Ebenso wurde auch der VOCs-Sensor nicht berücksichtigt, da bei diesem nicht nachvollziehbar ist, auf welche Gase er reagiert. Zusätzlich zu den Getränken wurden zur Klassifikation feste Lebensmittel einbezogen, sodass Daten von 24 Lebensmittelklassen aufgezeichnet und klassifiziert wurden (siehe Tabelle 8.2). Um einen Vergleich mit der Untersuchung von Hirano et al. (2015) zu ermöglichen, wurden ähnliche Getränke als Klassen verwendet, eine vergleichbare Datenaufzeichnung durchgeführt und der identische Klassifikator mit demselben Validierungsverfahren eingesetzt. Bis auf die Getränke „Moscato“ (Weißwein) und „Traubensaft“ (Fruchtsaft) wurden Daten von allen Getränken äquivalent zur Arbeit von Hirano et al. (2015) aufgezeichnet. Als Ersatz wurden Dornfelder (Rotwein) und Grapefruitsaft (Fruchtsaft) verwendet, um dieselbe Klassenanzahl zu erreichen.

Versuchsablauf

Abbildung 8.6 zeigt den Versuchsaufbau kurz vor einer Messung. Von jedem Getränk wurden eine Menge von 50 ml und von jedem festen Lebensmittel 10 g in ein Gefäß innerhalb eines abgedichteten Containers (l×b×h: 29×16×10,5 cm) platziert. In diesem befanden sich die beiden stationären Gas-Sensorboxen. Die zugehörigen Raspberry-Pi-3B-Einplatinencomputer wurden

Lebensmittel	Art
Apfel (apple)	Frucht
Banane (banana)	Frucht
Schwarzer Tee (black tea)	Heißgetränk
Buttermilch (buttermilk)	Milchprodukt
Karottensaft (carrot juice)	Gemüsesaft
Chardonney (chardonnay)	Weißwein
Kaffee (coffee)	Heißgetränk
Cola (cola)	Softdrink mit CO ₂
Dornfelder (dornfelder)	Rotwein
Emmenthaler (emmenthaler)	Käse
Knoblauch (garlic)	Gemüse
Grapefruitsaft (grapefruit juice)	Fruchtsaft
Grüner Tee (green tea)	Heißgetränk
Hirtenkäse (herder's cheese)	Käse
Heißes, ungechlortes Leitungswasser (hot tap water)	Wasser
Merlot (merlot)	Rotwein
Milch (milk)	Milchprodukt
Zwiebel (onion)	Gemüse
Orangensaft (orange juice)	Fruchtsaft
Birne (pear)	Frucht
Quark (quark)	Milchprodukt
Sekt (sparkling wine)	Weißwein mit CO ₂
Ungechlortes Leitungswasser (tap water)	Wasser
Tomatensaft (tomato juice)	Gemüsesaft

TABELLE 8.2: Lebensmittel, deren Gase gemessen wurden. Fett geschriebene Getränke sind vergleichbar mit denen aus (Hirano et al., 2015).

außerhalb des Containers platziert, da diese zusätzlich die Temperatur erhöht hätten. Nach dem Schließen des Containers wurden die Sensordaten für ca. 180 Sekunden aufgezeichnet. Nach jeder Messung wurde der Container geöffnet und für 10 Minuten mit einem Ventilator gelüftet, um die Gaskonzentrationen, Temperatur und Luftfeuchtigkeit „zurückzusetzen“. Der Raum wurde hierbei regelmäßig gelüftet. Nach einer ersten Messung aller Proben wurde eine weitere Messreihe durchgeführt, um kurzfristige Änderungen der Umgebung kompensieren zu können.

Ergebnisse

Der aus den Messungen gewonnene Datensatz wurde für die Klassifizierung auf eine Abtastrate von 1 Hz mittels Upsampling konvertiert. Die Zeitseriendaten wurden daraufhin in eine Fenstergröße von 5 Sekunden, ohne Überlappung, aufgeteilt. Für jedes Datenfenster wurden folgende Merkmale berechnet: Mittelwert, lineare Regression und Varianz. Für die Klassifikation wurden die Daten beider Messreihen kombiniert, wodurch 80 Fenster pro Lebensmittelklasse entstanden.

Zur Klassifikation wurde der C4.5 (J48) Entscheidungsbaum-Klassifikator von WEKA (3.9.2) (E. Frank et al., 2016) genutzt und eine zehnfache Kreuzvalidierung mit stabilem Zufallswert (random seed) bei allen Evaluationen eingesetzt.

Vergleich mit Hirano, Hayes und Truong (Getränke) Für diese Klassifizierung wurden nur die Daten der MQ-Sensoren und dem Luftfeuchtigkeitssensor genutzt, wobei drei Sensoren weniger als bei (Hirano et al., 2015) zur Verfügung standen. Es wurden ausschließlich die Daten der Getränke verwendet. Es ist zu beachten, dass die Sensoren nicht identisch waren und es zudem unmöglich war, identische Lebensmittel zu nutzen. Dennoch konnten vergleichbare Klassifikationsergebnisse

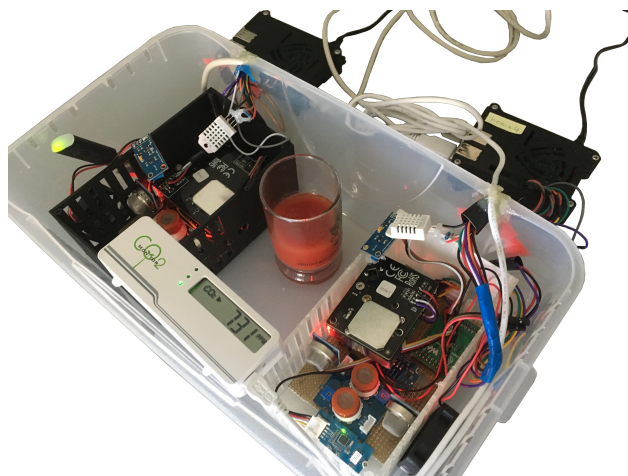


ABBILDUNG 8.6: Versuchsaufbau mit beiden stationären Gas-Sensorboxen und zu messendem Getränk im geöffneten Behälter (Foto von Chi Tai Dang, eigene Nachbearbeitung).

erreicht werden. Bei Hirano et al. (2015) betrugen die gewichtete mittlere Genauigkeit und Trefferquote 88,2 % bzw. 88,1 %, in der eigenen Untersuchung wurden 88,7 % bzw. 88,6 % erreicht. Die Konfusionsmatrix der eigenen Klassifikation ist in Abbildung 8.7 a) dargestellt.

a) MQ + humidity (recall: 88.6 % / precision: 88.7 %)															b) MICS + humidity (87.7 % / 87.7 %)															actual class
buttermilk	0	79	0	0	0	0	0	2	1	0	0	0	0	0	0	75	0	0	0	3	0	0	2	1	1	0	0	0	0	
chardonnay	1	0	54	0	4	0	0	20	0	0	0	0	3	0	0	1	2	57	0	7	4	0	3	2	0	0	0	6	0	0
cola	2	0	0	78	0	1	0	0	0	2	0	0	0	0	0	2	0	1	69	0	0	0	2	4	0	3	2	0	0	0
dornfelder	3	0	9	0	67	0	0	2	0	0	0	0	4	0	0	3	0	4	0	68	0	0	9	0	0	0	0	1	0	0
grapefruit j.	4	0	0	1	0	77	0	0	0	0	1	0	1	1	0	4	2	0	2	0	73	0	0	0	0	3	0	0	1	0
coffee	5	0	0	0	0	0	76	0	0	0	2	2	0	1	0	5	0	0	0	0	66	0	0	0	0	15	0	0	0	0
merlot	6	0	25	0	6	0	0	45	0	0	0	0	0	5	0	0	6	0	7	0	4	0	0	67	0	0	0	3	0	0
milk	7	5	0	0	0	0	0	0	75	1	0	0	0	0	0	7	0	1	1	0	0	0	0	79	0	0	0	0	0	0
carrot juice	8	0	0	0	3	0	0	0	0	1	77	0	0	0	0	8	0	0	0	0	0	0	0	0	80	1	0	0	0	0
orange juice	9	3	0	0	0	1	1	0	0	0	76	0	0	0	0	9	0	0	6	1	1	0	0	0	0	71	1	1	0	0
black tea	10	0	0	0	0	0	0	0	0	2	0	78	0	1	1	10	0	0	0	0	0	12	0	0	0	0	69	0	1	0
sparkl. wine	11	0	4	0	7	0	0	5	0	0	0	0	66	0	0	11	2	2	0	2	2	1	3	1	0	2	0	66	0	1
tap water	12	0	0	0	0	1	0	0	0	0	1	0	0	80	0	12	0	0	0	0	0	0	0	0	0	0	1	0	81	0
hot tap water	13	0	0	0	0	0	0	0	0	0	0	0	0	0	80	13	0	0	0	0	0	0	0	0	0	0	1	2	0	77
	0	1	2	3	4	5	6	7	8	9	10	11	12	13		0	1	2	3	4	5	6	7	8	9	10	11	12	13	
	prediction																													

ABBILDUNG 8.7: Klassifizierung von Getränken mit MQ-Sensoren bzw. MiCS-Sensor und Luftfeuchtigkeit. Prozentangaben: (Trefferquote / Genauigkeit).

Es ist erkennbar, dass die Weine „Chardonnay“ und „Merlot“ häufig miteinander verwechselt wurden. Dasselbe Problem trat seltener mit den beiden anderen alkoholischen Getränken Sekt und Dornfelder auf. In der Arbeit von Hirano et al. (2015) konnten ebenfalls bestimmte alkoholische Getränke schlechter unterschieden werden. Es handelte sich bei dieser allerdings um Chardonnay, Champagner und Merlot. Insofern war hier ein ähnliches Klassifizierungsverhalten erkennbar.

Vergleich MQ-Sensoren mit MiCS-Sensor (Getränke) Der modernere, kompaktere und stromsparendere MiCS-Sensor wurde in der eigenen Untersuchung mit den MQ-Sensoren verglichen. Die Luftfeuchtigkeit wurde hier ebenfalls einbezogen. Die Konfusionsmatrix ist in Abbildung 8.7 b) abgebildet. Die Genauigkeit und Trefferquote sind nahezu identisch 87,7 % und somit ca. 1 % geringer als bei den MQ-Sensoren. Die Klassifikation mit den MiCS-Sensordaten zeigte weniger Verwechslungen bei alkoholischen Getränken, jedoch wurden die beiden koffeinhaltigen Getränke (schwarzer Tee und Kaffee) häufiger verwechselt.

MQ-Sensoren, MiCS-Sensor, Luftfeuchtigkeit und CO₂ (alle Lebensmittel) Abschließend wird das Klassifikationsergebnis aller aufgezeichneten Sensoren vorgestellt (siehe Abbildung 8.10). Es wurden eine Genauigkeit und Trefferquote von 97,2 % bzw. 97,1 % erreicht. Im Vergleich zum besten Ergebnis der Kombination aus MQ-Sensoren mit Luftfeuchtigkeit und CO₂ war dies eine geringe Verbesserung von ca. 1 % bei Genauigkeit und Trefferquote. Aus praktischer Sicht stand die Verbesserung jedoch nicht in Relation zum zusätzlichen Stromverbrauch, den die MQ-Sensoren benötigten, auch wenn die Verbesserung gegenüber der besten Kombination mit MiCS-Sensor noch etwas höher ausfiel: 2,4 % für beide Metriken.

MQ + MiCS + humidity + CO₂ (recall: 97.1 % / precision: 97.2 %)

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	
apple	0	82	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	actual class
banana	1	0	81	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
pear	2	0	0	80	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	
buttermilk	3	0	0	0	79	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	0	
chardonnay	4	0	0	0	0	75	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
cola	5	2	0	0	0	0	78	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	
dornfelder	6	0	0	0	0	4	0	77	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	
emmentaler	7	0	0	0	0	0	0	81	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
grapefruit j.	8	0	1	0	0	0	0	0	79	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	
green tea	9	0	0	0	0	0	0	0	0	77	0	0	0	0	0	0	0	5	0	0	0	0	0	0	
herder's che.	10	0	0	0	0	0	0	0	0	0	82	0	0	0	0	0	0	0	0	0	0	0	0	0	
coffee	11	0	0	0	0	0	0	0	0	0	0	80	0	0	0	0	1	0	0	0	0	0	0	0	
garlic	12	0	0	0	3	0	0	0	0	0	0	0	79	0	0	0	0	0	0	0	0	0	0	0	
merlot	13	0	0	0	0	1	0	1	0	0	0	0	0	79	0	0	0	0	0	0	0	0	0	0	
milk	14	0	0	0	0	1	0	0	0	0	0	1	0	79	0	0	0	0	0	0	0	0	0	0	
carrot juice	15	0	0	0	0	0	0	0	0	0	0	0	0	0	78	0	3	0	0	0	0	0	0	0	
orange juice	16	0	0	0	0	0	0	4	0	0	0	0	0	0	1	76	0	0	0	0	0	0	0	0	
quark	17	0	0	0	2	0	3	0	0	0	0	0	0	0	1	0	75	0	0	0	0	0	0	1	
black tea	18	0	0	0	0	0	0	0	0	2	1	0	0	0	0	0	0	79	0	0	0	0	0	0	
sparkl. wine	19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	82	0	0	0	0	0	0	
tomato juice	20	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	81	0	0	0	0	0	
tap water	21	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	80	0	0	0	0	
hot tap water	22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	80	0	0	0	
onion	23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	82	0	
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	prediction

ABBILDUNG 8.10: Klassifizierung von Lebensmitteln mit allen Sensoren: MQ-Sensoren, MiCS-Sensor, Luftfeuchtigkeit und CO₂. Prozentangaben: (Trefferquote / Genauigkeit).

Zusammenfassung

Ziel der Arbeit war, die Ergebnisse von Hirano et al. (2015) zur Klassifizierung von Getränken (in geschlossenen Behältern) nachzuvollziehen und auf Speisen zu erweitern. Zusätzlich wurde als Ersatz für die MQ-Sensoren die Nutzung eines moderneren, kompakteren und stromsparenderen MiCS-Sensors geprüft. Eine Verbesserung der Erkennungsraten wurde unter Nutzung verschiedener Sensorkombinationen (MQ-, MiCS-, CO₂-, Luftfeuchtigkeitssensoren) untersucht.

In der eigenen Arbeit konnte die Klassifizierung von Getränken in geschlossenen Behältern mit MQ-Sensoren nachvollzogen werden. Eine zusätzliche Erkennung von Speisen war ebenfalls möglich. Die mit dem MEMS-Verfahren produzierten MiCS-Sensoren konnten die MQ-Sensoren mit nahezu identischen Erkennungsraten ersetzen. Eine Kombination aller Sensoren erreichte die höchste Erkennungsrate.

Bei den in dieser Arbeit vorgestellten Ergebnissen handelte es sich ausschließlich um Messungen, die in weitestgehend kontrollierter Umgebung (Container) stattfanden, wodurch störende Einflüsse reduziert wurden, die bei realistischen Messungen z. B. in einer Küche auftreten würden. Untersuchungen von Hirano et al. (2015) mit einer Dunstabzugshaube zeigten, dass die Klassifizierung von drei Getränken in einem Raum mit entsprechender Ventilation grundsätzlich möglich ist. Es befand sich allerdings immer nur ein Getränk im Raum. Das Auftreten isolierter Gerüche ist unter normalen Bedingungen allerdings eher die Ausnahme. Ein spezielles kompaktes mobiles Sensorgerät mit MEMS-Gassensorik und Lüfter könnte allerdings eingesetzt werden, um beispielsweise Getränke oder bestimmte Gerichte komfortabler erfassen zu können.

In Abschnitt 3.5.3 wurden bereits generelle Probleme mit MOS- und teils NDIR-basierten Gassensoren vorgestellt, die auch deren Anwendung bei Klassifizierungen stark einschränken, da hierdurch trainierte Modellen kaum weiterverwendet werden können. In dieser Arbeit wird im

Folgenden nur auf eine Problemlösung für das einzigartige Sensorverhalten ausführlicher eingegangen. Allerdings findet sich auch zu weiteren Problemen, wie beispielsweise dem kontinuierlichen Sensordrift, weiterführende Literatur mit Lösungsansätzen z. B. von Ma et al. (2018).

8.3.3 Erkennung von Lebensmitteln mittels Gassensorik: Transferlernen zur Kompensation des einzigartigen Sensorverhaltens

Einleitung

In diesem Abschnitt wird eine Lösung für das Problem des einzigartigen Sensorverhaltens (siehe Abschnitt 3.5.3) von MOS-Sensoren vorgestellt, das auf dem Transferlernen-Verfahren von Laref et al. (2018) basiert. Produktionsbedingt unterscheiden sich die Signale von MOS-Sensoren desselben Typs bezüglich ihres Grundwerts und dem Wertebereich (Empfindlichkeit) sehr deutlich. Aus diesem Grund entwickelten Laref et al. (2018) ein Verfahren, um eine Kalibrierung eines Gassensors zur Messung von Luftverschmutzung mit einem anderen Sensor desselben Typs nutzen zu können. Hierdurch sollte eine deutlich aufwändigere und teurere Neukalibrierung des anderen Sensors verhindert werden. Das einzigartige Sensorverhalten von MOS-Gassensoren verhindert auch, dass Klassifizierungsmodelle mit anderen Gassensoren desselben Typs verwendet werden können. Das im vorherigen Abschnitt mit den Daten der einen Gas-Sensorbox trainierte Klassifizierungsmodell ist beispielsweise nicht mit der anderen Sensorbox nutzbar. In der eigenen Arbeit wurde daher untersucht, ob das Verfahren von Laref et al. (2018) durch entsprechende Anpassungen auch für Klassifizierungsprobleme eingesetzt werden kann.

Für diese Untersuchung wurden die synchronisierten Daten der beiden stationären Gas-Sensorboxen verwendet, die für die im vorherigen Abschnitt beschriebene Nahrungsmittel-Klassifizierung aufgezeichnet wurden. In diesem Abschnitt wurde ein SVM-Klassifikator genutzt und die Daten mussten für das Training und die Evaluation aufgeteilt werden (keine Kreuzvalidierung), daher sind die Evaluationsergebnisse nicht direkt mit denen des vorherigen Abschnitts vergleichbar. Teile dieses Abschnitts wurden in der eigenen Arbeit (Seiderer, Dang & André, 2019) publiziert.

Verfahren

Bei dem von Laref et al. (2018) vorgestellten Verfahren handelt es sich um eine Art des Transferlernens. Es kann folgendermaßen kategorisiert werden (siehe Abschnitt 5.3 im Kapitel „Transferlernen“). Es wird eine Standardisierung der Sensordaten durchgeführt, die es ermöglicht, das Modell mit anderen Sensoren desselben Typs zu nutzen. Die Domänen sind identisch, da dieselben Merkmale genutzt werden und die Randverteilungen gleich sind (identische Klassen, identische Datenmenge pro Klasse und Sensor): $\mathcal{D}_S = \mathcal{D}_T$, da $\mathcal{X}_S = \mathcal{X}_T$ und $P(X_S) = P(X_T)$. Es unterscheidet sich jedoch die Quell- und Zielaufgabe, dadurch, dass die Sensoren ähnlich, jedoch nicht identisch sind und somit die Funktionen zum Vorhersagen der Beschriftungen verschieden sind: $\mathcal{T}_S \neq \mathcal{T}_T$, da zwar $\mathcal{Y}_S = \mathcal{Y}_T$ aber $f_S(\cdot) \neq f_T(\cdot)$. Beschriftungen der Daten sind in der Quell- und Zieldomäne vorhanden: D_S und D_T . Es handelt sich daher um überwacht (induktives) maschinelles Lernen. Die Quell- und Zielaufgabe werden jedoch nicht gleichzeitig gelernt. Es findet ein Wissenstransfer von Merkmalsrepräsentationen statt, die ein Regressionsmodell lernt und anwendet.

In dieser Untersuchung lag der Fokus nicht auf der Reduktion benötigter Daten, daher wurde im Vergleich zur Arbeit von Laref et al. (2018) darauf verzichtet, den SPXY-Kernel zur Selektion der Daten für die Standardisierung zu verwenden. Als letzter Schritt wurde hier eine SVM zur Klassifizierung der Lebensmittel genutzt, was im ursprünglichen Verfahren nicht der Fall war. Die c- und gamma-Parameter für die SVM-Klassifikation wurden im Bereich von $[10^{-7}; 10^7]$ in 10er-Schritten (225 Kombinationen) automatisch getestet.

Das verwendete Verfahren ist in Abbildung 8.11 dargestellt. Die Evaluationen **a)**, **b)** und **c)** sind in Abbildung 8.12 in der oberen, mittleren und unteren Reihe aufzufinden. Eine der Sensorboxen wurde als „master“, die andere als „slave“ bezeichnet. Zunächst wurden die Daten

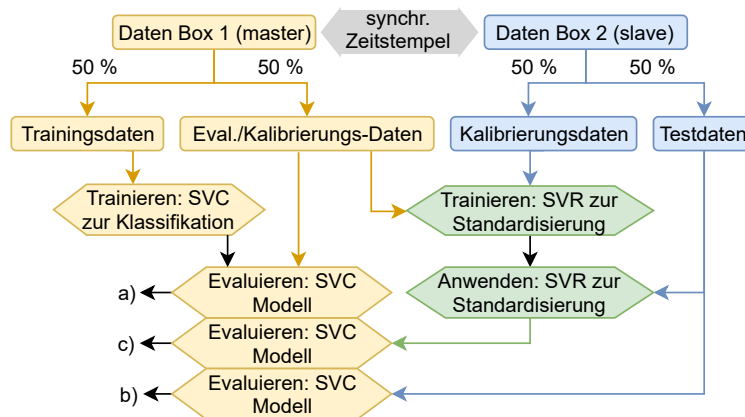


ABBILDUNG 8.11: Verwendetes Verfahren zur Übertragung der Kalibrierung. Konfusionsmatrizen zu a), b), c) siehe Abbildung 8.12 obere, mittlere bzw. untere Reihe.

mittels Upsampling auf 1 Hz konvertiert und anschließend mittels ihrer Zeitstempel synchronisiert. Hierdurch waren ca. 400 Datenwerte pro Sensorbox, Sensor und Klasse verfügbar. Die Hälfte der Daten der Master-Box wurden für das SVM-Klassifizierungsmodell (SVC) zum Trainieren genutzt, die andere Hälfte diente der Evaluation dieses Modells (a) siehe Abbildung 8.12 oben) bzw. dem Training der SVM-Regression (SVR). Die Daten der Slave-Box wurden ebenfalls halbiert und als Kalibrierungsdaten für die SVR genutzt. Auf diese Werte sollten die entsprechenden Daten der Master-Box abgebildet werden. Der restliche Datensatz der Slave-Box wurde zum Testen verwendet, um die Standardisierung mit der SVR auf diesen Daten anzuwenden (c) siehe Abbildung 8.12 unten) und um zu untersuchen, wie sich das Klassifizierungsmodell verhält, wenn die Daten nicht standardisiert wurden (b) siehe Abbildung 8.12 mitte).

Für die Tests wurden die Rohdaten des MiCS-Sensors (aller Sensorelemente: RED, OX, NH₃) bzw. die des MQ2-Sensors verwendet. Der MQ2-Sensor zeigte beim Transfer der Kalibrierung die besten Ergebnisse, daher werden die Ergebnisse der anderen MQ-Sensoren hier nicht aufgeführt. Die Parameter des SVM-Klassifikators (SVC) wurden für den MiCS-Sensor hinsichtlich Genauigkeit optimiert, da hier auch eine hohe Trefferquote zu beobachten war. Beim MQ2-Sensor wurde das Maximum der Summe von Genauigkeit und Trefferquote verwendet. Die Parameter-Optimierungen wurden für jede Klassifikation neu durchgeführt.

Wie zu erwarten, waren die Klassifikationsergebnisse am besten, wenn die Test-Daten von derselben Box stammten, mit denen das Modell trainiert wurde. Die Genauigkeit und Trefferquote waren nur beim MQ2-Sensor generell geringer, da dieser weniger Gaskonzentrationen erfassen kann. Sobald nicht-standardisierte Test-Daten der anderen Box (slave) verwendet wurden, brachen die Erkennungsraten stark ein. Nach der Standardisierung waren die Daten der Slave-Box wesentlich besser verwendbar mit dem Modell der Master-Box. Allerdings waren die Genauigkeiten und Trefferquoten nahezu um die Hälfte reduziert als, wenn Sensordaten nur von derselben Box genutzt wurden. Im Vergleich zu den nicht-standardisierten Sensordaten war jedoch eine deutliche Erhöhung der Metriken erkennbar.

Zusammenfassung

Die Untersuchung zeigt, dass mit dem in der eigenen Arbeit für Klassifizierungen angepassten und erweiterten Verfahren von Laref et al. (2018) eine deutliche Erhöhung der Erkennungsraten, trotz der Nutzung eines anderen MOS-Gassensors desselben Typs, erreicht werden kann. In zukünftigen Untersuchungen könnten weitere Techniken untersucht werden, um die Reduktion der Erkennungsleistung mit Daten baugleicher Gassensoren im Vergleich zum ursprünglich trainierten Modell weiter zu verringern, sodass mit weniger Aufwand allgemeingültige Klassifizierungsmodelle für verschiedene Anwendungszwecke erstellt werden können.

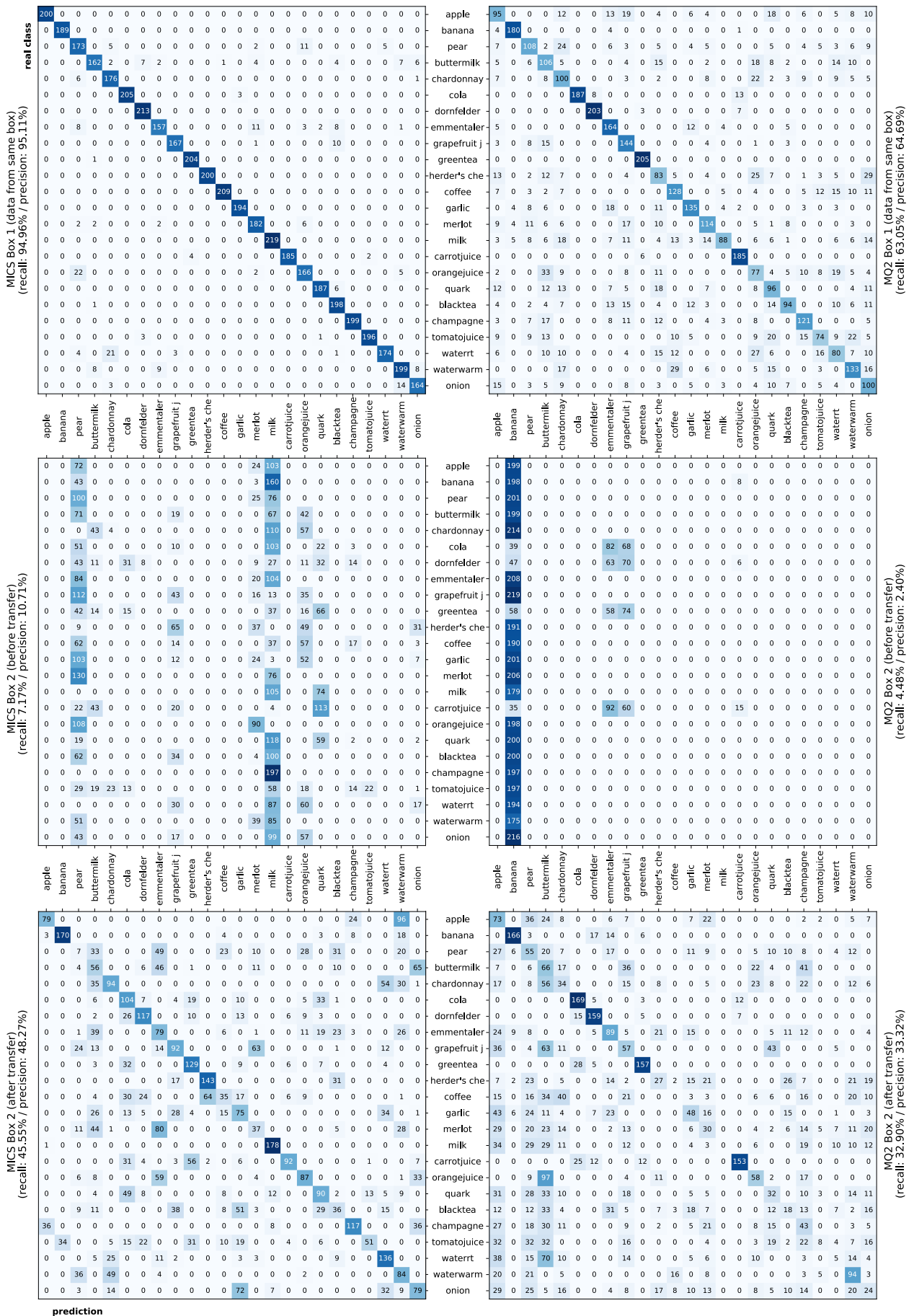


ABBILDUNG 8.12: Ein Modell wurde mit den Sensordaten der Box 1 (master) trainiert. Links wurden ungenutzte Daten der Box 1 zur Evaluation genutzt. In der Mitte Sensordaten der Box 2 (slave). Rechts wurden die Sensordaten von Box 2 zuvor standardisiert. In der oberen Reihe wurden Daten des MiCS-Sensors verwendet in der unteren Reihe vom MQ2-Sensor. (waterrt = water room temperature)

8.3.4 Vereinfachung der mobilen Ernährungserfassung mittels Interaktion mit Smart-scale, Smartwatch und Smartphone

Einleitung

Die Protokollierung der exakten Art und Menge verzehrter Nahrungsmittel stellt eine zeitaufwändige Aufgabe dar und wird daher von vielen Personen nach kurzer Zeit eingestellt, da ihnen der Aufwand zu hoch ist (Cordeiro, Epstein et al., 2015). Um diesen bei der Nahrungsprotokollierung zu reduzieren, wurde ein Multi-Geräte-System bestehend aus einem Smartphone, einer Smartwatch und einer Smartscale entwickelt und evaluiert, das in diesem Abschnitt vorgestellt wird. Die Untersuchung wurde in den eigenen Publikationen (Seiderer & André, 2017; Seiderer, Flutura & André, 2017) veröffentlicht.

Die einzelnen mobilen Geräte besitzen unterschiedliche Eigenschaften, die von einem mobilen System zur Nahrungserfassung vorteilhaft kombiniert werden sollten. Das Smartphone ermöglicht hierbei die flexibelsten Nutzereingaben und reduziert bereits signifikant den Aufwand im Vergleich zum traditionellen Aufschreiben mit Stift und Papier (Y. Kim et al., 2016). Eine Smartwatch kann ein Smartphone ergänzen, da diese immer in Reichweite ist und unauffälliger verwendet werden kann. Durch den kleinen Bildschirm sind allerdings die Eingabemöglichkeiten beschränkt. Eine Smartwatch ist insbesondere in Situationen hilfreich, in denen ein Nutzer in Eile ist und Nahrungsmittel erfasst werden sollen, die bereits aus Listen (Favoriten oder kürzlich erfasste Lebensmittel) ausgewählt werden können. Ein Protokolleintrag mit fehlenden Daten, wie beispielsweise der Menge, ist bereits von großem Nutzen, da der Anwender die fehlenden Informationen mit dem Smartphone später ergänzen kann. Selbst unvollständige Einträge können jedoch medizinisch relevante Informationen über die Uhrzeit und Häufigkeit der Nahrungsaufnahme besitzen (Y. Kim et al., 2016).

In der Regel ist es bei der Ernährungsprotokollierung notwendig, die Menge von Essen und Trinken zu erfassen. Eine Mengenabschätzung ist allerdings in vielen Fällen schwierig und somit ungenau. Daher war bei diesem System zusätzlich ein Smartscale-Prototyp enthalten, der die Gewichtsdaten per Bluetooth übertragen konnte. Für eine komfortable Nutzung wurde bei der Entwicklung darauf geachtet, dass der Prototyp möglichst einfach mit den Nutzergeräten kommunizieren konnte, eine lange Akkulaufzeit besaß sowie möglichst kompakt war. Die Waage konnte das aktuell gewogene Gewicht per Bluetooth parallel an das Smartphone und die Smartwatch senden. Die war zudem für mehrere Wochen ohne erneute Aufladung des Akkus funktionsfähig. Zusätzlich wurde die Waage sehr kompakt gestaltet, sodass es möglich war, sie überall zu platzieren, wo der Nutzer sie verwenden wollte. Um eine hohe Flexibilität bei der Nutzung des Systems zu gewährleisten, war die Verwendung der Smartwatch sowie der Smartscale optional.

Die Nutzung mehrerer Geräte für die Nahrungsmittelerfassung scheint eine vielversprechende Option darzustellen, allerdings ist das Design der Interfaces für mehrere Geräte keine triviale Aufgabe, da ein fließender Übergang zwischen diesen ermöglicht werden sollte.

Zum Zeitpunkt der Arbeit war diese Systemkombination, die eine Smartwatch und eine Smartscale mit einem Smartphone für die Erfassung von Nahrungsmitteln verbindet, neuartig. In der Literatur sind jedoch bezüglich der Multi-Geräte-Interaktion folgende wesentliche Vorarbeiten aufzufinden, die bei der Entwicklung des eigenen Systems berücksichtigt wurden. In der Arbeit von Dong et al. (2016) wurden zwei Dimensionen der Zwischengeräte-Usability (engl. inter-device usability) identifiziert. Nutzer sollten befähigt werden, Wissen, das sie während der Nutzung eines Gerätes gewonnen haben, auf ein anderes Gerät zu übertragen (Wissens-Kontinuität; engl. knowledge continuity). Zudem sollten sie eine Aufgabe, die auf einem Gerät begonnen wurde, auf einem anderen Gerät fortsetzen können (Aufgaben-Kontinuität; engl. task continuity). Beide Arten der Kontinuität könnten durch ein konsistentes Design der Nutzerinterfaces unterstützt werden, um es dem Nutzer zu ermöglichen, problemlos zwischen den Geräten zu wechseln. Ein weiterer Aspekt, der beim Design von Multi-Geräten-Interfaces berücksichtigt werden sollte, ist der zeitliche Ablauf der Aufgaben, die vom Nutzer durchgeführt werden. Sørensen et al. (2014)

unterschieden zwischen der simultanen und sequenziellen Ausführung von Aufgaben, die über mehrere Geräte verteilt sind. Bei der Nahrungserfassung könnte der Nutzer ein Lebensmittel auf der Smartscale platzieren, während die Art des Nahrungsmittels vom Nutzer per Smartphone ergänzt wird (simultane Nutzung). Zudem könnte ein Anwender mit dem Smartphone einen Eintrag vervollständigen, der zuvor mit der Smartwatch unterwegs angelegt wurde (sequenzielle Nutzung). In beiden Fällen müssen die automatisch oder manuell eingegebenen Informationen über das Lebensmittel zwischen den Geräten synchronisiert werden.

Mit dem entwickelten Prototypen sollten folgende Fragen bezüglich der Multi-Geräte-Interaktion beantwortet werden:

- Werden mehrere Geräte von Nutzern verwendet, wenn ja, wie werden sie eingesetzt?
- Welche Faktoren spielen bei der Gerätenutzung eine Rolle?
- Wann wird die Nahrungsmittelerfassung durchgeführt?

Im folgenden Abschnitt wird verwandte Literatur zu Erkennungstechniken und Nutzerinterfaces diskutiert, die (semi-)automatische Nahrungserfassung ermöglichen. Daraufhin werden die Anforderungen für ein Nahrungserfassungssystem auf Basis von mehreren Geräten identifiziert. Anschließend folgt eine Beschreibung, wie diese Anforderungen durch das Design und die Implementierung des Systems berücksichtigt wurden. Um den Multi-Geräte-Ansatz für die Lebensmittelerfassung zu evaluieren, wurde eine Feldstudie durchgeführt, deren Ergebnisse diskutiert werden.

Verwandte Arbeiten

Nahrungserfassung Ein Überblick über aktuelle Verfahren zur automatischen Nahrungserfassung ist am Anfang dieses Kapitels bereits aufgezeigt worden. Der in dieser Arbeit vorgestellte Prototyp wird von einer Smartscale unterstützt, ergänzende Automatisierungen, z. B. mittels Bilderkennung, wären möglich. Falls höhere Genauigkeit und Präzision benötigt werden, sind allerdings derzeit immer noch manuelle oder partiell automatisierte Systeme zur Erfassung der vollständigen Nahrungsaufnahme notwendig. Teilweise kann der Einbezug des Nutzers in den Erfassungsprozess als Vorteil angesehen werden, da dieser sich eher über sein Ess- und Trinkverhalten bewusst wird und mögliches Fehlverhalten hierbei bemerkt (Cordeiro, Bales et al., 2015; Y. Kim et al., 2016).

Smartphone- und Smartwatch-Apps Mobile Geräte bieten verschiedene Möglichkeiten, den Erfassungsvorgang zu vereinfachen und sind nahezu immer in Reichweite. In der Studie von Carter et al. (2013) wurde eine Smartphone-App gegenüber einer Webseite oder einem Tagebuch aus Papier bevorzugt. Nur die Nutzer der Smartphone-App zeichneten ihre täglich konsumierten Nahrungsmittel nach sechs Monaten noch auf.

Um einen Eindruck zu erhalten, welche Funktionen beliebte Smartphone-Apps anbieten, kann der Überblick von Franco et al. (2016) bezüglich ernährungsbezogener, mobiler Anwendungen herangezogen werden. Zusätzlich ist der Überblick über gesundheitsbezogene Apps von Murnane et al. (2015) hierbei durchaus hilfreich.

Eine App, die versucht die Ernährungserfassung durch einen multimodalen Ansatz zu vereinfachen, wurde von Lim et al. (2017) beschrieben. Es wurde hierzu Bilderkennung oder alternativ Texteingabe angeboten. Bei einer diesbezüglichen Untersuchung mit Studienteilnehmern zeigte sich, dass die manuelle Texteingabe trotz der bequemen Fotoaufnahmen bevorzugt wurde, da sie zuverlässiger funktionierte. Dies belegt, dass so lange automatische Systeme nicht nahezu ähnliche Genauigkeiten wie eine manuelle Eingabe erlauben, zusätzliche Eingabemethoden zur Korrektur oder als Alternative angeboten werden sollten.

Da Smartwatches eine geringere Verbreitung besitzen, existieren vergleichsweise wenige Apps und verwandte Literatur, die sich mit der Nahrungserfassung beschäftigen. Årsand et al. (2015) stellten eine Smartphone- und Smartwatch- (Pebble) App für Diabetes-Patienten vor, die als Tagebuch diene. Es konnte festgestellt werden, dass die Smartwatch den Eingabevorgang vereinfachte. Daher sollte dies auch bei einer generellen Nahrungserfassung der Fall sein.

Lutze et al. (2015) stellten eine spezielle Smartwatch-App für Senioren vor, die eine Dehydrierung verhindern und in Notfallsituationen Unterstützung bieten sollte. In diesem Fall wurde die Smartwatch als einziges Gerät eingesetzt. Sie zeichnete Sensordaten auf, gab Erinnerungen hinsichtlich „Trinken“ aus und ermöglichte es, dem Nutzer einfache Fragen zu beantworten, falls eine mögliche Notfallsituation erkannt wurde, bevor Hilfe gerufen wurde. In der Arbeit wurde offensichtlich, dass die Smartwatch den großen Vorteil besitzt, schnelle Eingaben zu ermöglichen und immer in Reichweite ist, wenn diese für eine Interaktion benötigt wird.

Gemäß der behandelten Literatur ist die Nutzung von Smartphones grundsätzlich vorteilhaft, da sie weit verbreitet und meistens in relativ kurzer Reichweite sind. Smartwatches zeigen zusätzliche Vorteile in Situationen, wenn eine sehr schnelle Eingabe notwendig ist, da mit diesen direkt interagiert werden kann.

Smartscale Im Nachfolgenden wird als Smartscale eine Waage bezeichnet, die direkt mit einem Smartphone oder einer Smartwatch kommunizieren kann, um eine einfache Übertragung des gewogenen Gewichts von Objekten zu ermöglichen.

SITU⁴ war eine kommerzielle Smartscale mit einer iOS-App, die Kalorien und Nährstoffe von Lebensmitteln mithilfe des Gewichts schätzen konnte, nachdem die Art des Nahrungsmittels eingegeben wurde. Im Vergleich zur SITU-Smartscale ist der in der eigenen Arbeit vorgestellte Prototyp kompakter und musste nicht mit einem Smartphone per Bluetooth gekoppelt werden, wodurch die Handhabung des Prototyps erleichtert wurde. Durch die Entwicklung eigener Software war es zudem leichter möglich, die Smartscale in eigene Anwendungen zu integrieren.

Einige spezielle Smartphones konnten ohne zusätzliche Geräte verwendet werden, um kleinere Objekte auf ihrem Bildschirm zu wiegen⁵. Obwohl dies eine interessante Option dargestellt hätte, war diese Funktion sehr gerätespezifisch, nicht so präzise wie eine Waage, funktionierte nur mit kleinen Lebensmitteln bzw. Gefäßen und das maximal messbare Gewicht war sehr gering. Daher wurden in der eigenen Arbeit Smartscales bevorzugt, die zum Wiegen von Lebensmitteln konstruiert wurden und im Vergleich zu automatischen Methoden oder Schätzungen des Gewichts eine hohe Genauigkeit besitzen.

Lessel et al. (2016) verwendeten eine Smartscale als Trinksensor für eine App, die mittels Gamifizierungen Nutzer dazu animieren sollte, mehr zu trinken. Soubam et al. (2017) präsentierten ein System, das eine Smartwatch beinhaltet, um eine Trinkaktivität zu erkennen und eine Smartscale, um die Flüssigkeitsmenge zu bestimmen. Die in den beiden Arbeiten präsentierten Prototypen zeigten einige Ähnlichkeiten zu dem, der im Rahmen der eigenen Arbeit entwickelt wurde. Dies liegt daran, dass es bei der Verwendung einer Wägezelle als Sensor nur begrenzte Designmöglichkeiten bei der Grundkonstruktion gibt. Unterschiede bestanden dahingehend, dass der Prototyp von Lessel et al. (2016) keine größeren Gefäße zur Messung aufnehmen konnte, da dies der obere Gehäuserand verhindert. Zusätzlich war es notwendig, einen Knopf an der Waage zu drücken, um das Gewicht an die Smartphone-App zu senden. Dieser Knopf hätte von einem größeren Teller verdeckt werden können. Der Smartscale von Soubam et al. (2017) fehlte ein Gehäuse. Der Arbeit ist zudem nicht zu entnehmen, ob eine Batterie oder Ladeschaltung integriert waren.

Ein wesentlicher Unterschied zu dem in der eigenen Arbeit vorgestellten Prototypen bestand in der Kommunikationsmethode zwischen der Smartscale und einem Smartphone. Diese wurde in der eigenen Arbeit mittels Bluetooth Low Energy (BLE) Advertisement-Paketen umgesetzt.

⁴<https://www.kickstarter.com/projects/situ/situ-smart-food-nutrition-scale?lang=de> (21.10.2019)

⁵<http://money.cnn.com/2015/09/03/technology/force-touch-weigh-objects-huawei/> (21.10.2019)

Hierdurch war es nicht nötig, die Geräte miteinander zu koppeln. Durch die Nutzung dieser Kommunikationsmethode und Techniken, die den Stromverbrauch deutlich verringerten, betrug die Batterielaufzeit des Prototyps im Normalfall ca. drei Wochen.

Anforderungen

In diesem Abschnitt werden die wesentlichsten Anforderungen an ein Nahrungserfassungssystem vorgestellt. Hierfür wurden Literaturbeiträge zur grundsätzlichen Problematik der Lebensmittelerfassung ausgewertet. Zudem wurden eigene semistrukturierte Interviews durchgeführt und beliebte Android-Apps untersucht. Diese Anforderungen wurden während der Entwicklung des Prototyps berücksichtigt. In einem späteren Abschnitt wird diskutiert, inwieweit diese im Prototypen umgesetzt wurden.

Literatur Viele Anforderungen konnten mittels der Arbeit von Cordeiro, Epstein et al. (2015) identifiziert werden, in der häufig auftretende Probleme bei der Lebensmittelerfassung gesammelt wurden. Die Autoren werteten hierzu die Fragebögen von 141 Personen aus, die bereits Erfahrungen mit Lebensmittelerfassung besaßen. Zusätzlich wurden von den Autoren Probleme aus Internetforeneinträgen von drei mobilen Lebensmitteltagebücher-Apps (engl. food journaling) ergänzt.

Ferner wurden zur Identifikation von Anforderungen für den eigenen Prototyp die Überblicke über ernährungsbezogene mobile Anwendungen von Franco et al. (2016) und über Gesundheits-Apps von Murnane et al. (2015) berücksichtigt.

Semistrukturierte Interviews Es wurden semistrukturierte Interviews mit 24 Personen (13 Frauen, 11 Männer) mit einem Alter zwischen 16 und 55 (Mittelwert: 26,2) Jahren, durchgeführt. Die meisten Teilnehmer waren Schüler oder Studierende. Alle Probanden waren mit Smartphones vertraut. Die Interviews beinhalteten ca. 25 Fragen hinsichtlich der Demografie, Smartphone-Nutzung, Erfahrung mit Gesundheits-Apps, Bereitschaft Smartwatches zu nutzen, Bedenken gegenüber der Privatsphäre, Kenntnis über Hilfsmittel zur Nahrungserfassung, Motivation Lebensmittel zu erfassen und hinsichtlich des bevorzugten Ablaufs bei der Erfassung. Einige der Fragen wurden für die Evaluation des Prototyps wiederverwendet.

Beliebte Apps Vier Smartphone-Apps wurden aus dem Android-App-Store nach ihrer Beliebtheit und Funktionalität ausgewählt, um zusätzliche Anforderungen zu identifizieren. „MyFitnessPal“ (2017: 1.469.230 Bewertungen), „S-Health“ (2017: 233.598 Bewertungen), „FatSecret“ (2017: 171.922 Bewertungen) und „Lifesum“ (2017: 70.224 Bewertungen). Die meisten Anwendungen waren nicht hinsichtlich einer möglichst effizienten Nahrungserfassung optimiert, sondern beinhalteten zusätzliche Funktionen wie z. B. eine Statistik, gesundheitsbezogene Informationen und das Aufzeichnen von Nutzeraktivitäten. Für die weiteren Untersuchungen wurden allerdings ausschließlich die Funktionalitäten bei der Lebensmittelerfassung berücksichtigt. Während der Entwicklungszeit des Prototyps waren keine beliebten Apps mit Smartwatch-Integration zur Aufzeichnung von Essen auffindbar. Ausschließlich S-Health ermöglichte auf der Samsung Gear S2 Smartwatch eine Eingabe von Getränken mit der Anzahl an Gläsern und Tassen.

Die einzelnen Apps wurden für jeweils einen Tag zur Aufzeichnung der Nahrungsaufnahme eingesetzt. Der Funktionsumfang wurde im Detail untersucht. Zu den Funktionen zählten beispielsweise die Erkennung von Barcodes zur Eingabe von Lebensmitteln, Aktivitätserfassung und Details über gesunde Ernährung (z. B. empfohlene Aufnahme von Vitaminen). Ansätze für Gamifizierungen und Visualisierung wurden gesammelt, die für künftige Prototypen verwendbar wären.

Reduktion des Aufwands bei der Lebensmittelerfassung Die Hauptaufgabe des zu entwickelnden Prototypen war, den Aufwand bei der Nahrungsaufnahme zu reduzieren. Mehrere Anforderungen, die zur Erreichung dieses Ziels berücksichtigt werden sollten, wurden identifiziert. Die Quelle der jeweiligen Anforderung ist in Klammern angegeben.

- Der Erfassungsvorgang sollte so einfach wie möglich sein, da hoher Aufwand eine große Hürde darstellt (Cordeiro, Epstein et al., 2015).
- Das System sollte bei der Ermittlung der Menge / Masse einer Portion behilflich sein, da dies häufig ein Problem ist (Cordeiro, Epstein et al., 2015).
- Eine Datenbank mit Lebensmitteln sollte verwendet werden, um die Texteingaben zu vereinfachen (beliebte Apps).
- Nutzerfavoriten sollten möglich sein (beliebte Apps).
- Der Nutzer sollte aus einem Verlauf von kürzlich eingetragenen Lebensmitteln auswählen können (beliebte Apps).
- In weniger öffentlichen Situationen könnte Spracherkennung als alternative Eingabemethode angeboten werden (Interviews).

Privatsphäre, Flexibilität der Nutzung und Sonstige Hinsichtlich Privatsphäre, Flexibilität der Nutzung, Erinnerungen und Möglichkeiten zur Selbstreflexion, wurden weitere Anforderungen identifiziert.

- Das Eintragen sollte in einer unauffälligen Weise möglich sein, da manche Personen nicht möchten, dass z. B. Freunde oder Kollegen dies bemerken (Cordeiro, Epstein et al., 2015).
- Das System sollte eigene Notizen zulassen (Interviews).
- Nutzer sollten Einträge editieren können (Interviews).
- Erinnerungen an die erforderliche Erfassung sollten gegeben werden, da Personen dazu neigen, das Eintragen zu vergessen (Y. Kim et al., 2016).
- Nutzer sollten über ihre Nahrungsaufnahme reflektieren können (Cordeiro, Epstein et al., 2015).

Multi-Geräte-Interaktion Der Aufwand bei der Lebensmittelerfassung sollte durch die Kombination mehrerer Geräte mit ihren speziellen Vorteilen reduziert werden. Die nachfolgenden Anforderungen wurden berücksichtigt, um eine hohe Usability zwischen den Geräten zu ermöglichen.

- Damit ein Nutzer problemlos zwischen den Geräten wechseln kann, wird eine hohe Konsistenz bei den eingesetzten GUIs benötigt, wobei gleichzeitig die UI-Standards der entsprechenden Plattform (z. B. Android) berücksichtigt werden müssen (Dong et al., 2016).
- Datensynchronisation muss zwischen allen Geräten unterstützt werden, um eine synergetische Nutzung zwischen den Geräten zu ermöglichen (Sørensen et al., 2014).

Die Funktionalität des Gesamtsystems sollte zwischen den verschiedenen Geräten aufgeteilt werden, um deren spezifische Stärken nutzen zu können (Dong et al., 2016). Mit dem Smartphone konnten bei diesem System alle Funktionen genutzt werden, da es hinsichtlich der Eingabemöglichkeiten und der Hardware das leistungsfähigste Gerät darstellte. Insbesondere in Situationen, in denen eine schnelle oder verdeckte, private Lebensmittelerfassung vom Nutzer erwünscht wurde, war die jederzeit bedienbare Smartwatch besser geeignet als das Smartphone. Durch ihr sehr kompaktes Touch-Display und die somit vereinfachte grafische Oberfläche, verfügte sie allerdings

nur über die relevantesten Funktionen zur Erfassung. Die Smartscale ergänzte das Smartphone und die Smartwatch bezüglich ihrer Fähigkeit, die Menge eines Lebensmittels komfortabel und präzise wiegen zu können. Sie konnte optional während des Erfassungsprozesses mit dem Smartphone oder der Smartwatch zur automatischen Übertragung des Lebensmittelgewichts eingesetzt werden.

Smartscale-Prototyp Der Smartscale-Prototyp sollte folgende Anforderungen erfüllen:

- Die Waage / der Prototyp sollte leicht zu transportieren sein.
- Ihr Messbereich und ihre Stabilität sollten das Wiegen eines Lebensmittels, ggf. auch mit Gefäß, bis zu 5 kg ermöglichen.
- Die Messgenauigkeit sollte eine hinreichende Präzision aufweisen.
- Der Prototyp sollte einen relativ geringen Preis besitzen.

Diese Anforderungen waren vergleichbar mit denen von Lessel et al. (2016), sie wurden für den eigenen Prototyp jedoch erweitert:

- Der Akkuverbrauch sollte gering sein (weniger Ladeaufwand).
- Verzicht auf die Verwendung von Tastern zum Wiegen.
- Verzicht auf eine Bluetooth-Kopplung.
- Es sollte möglich sein, Teller oder größere Gefäße auf die Waage zu stellen.

Design der Smartscale

Entwicklungsprozess Um das Gewicht von Objekten zu wiegen, werden meistens resistive Drucksensoren oder Wägezellen eingesetzt. Ein **erster Prototyp** nutzte vier resistive Drucksensoren (siehe Abbildung 8.13). Die Auswertung der Messergebnisse zeigte, dass das Datensignal relativ stark mit Rauschen versehen und zudem eine stetig zunehmende Sensordrift vorhanden war. Präzise Gewichtsdaten können aus solchen Daten nur mit hohem Verarbeitungsaufwand gewonnen werden. Die Möglichkeit einer Verwendung von resistiven Drucksensoren wurde somit nicht weiter verfolgt.



ABBILDUNG 8.13: Bilder vom ersten Smartscale-Prototyp (resistive Drucksensoren).

Der **zweite und dritte Prototyp** nutzten daher Wägezellen, die meistens auch in Küchenwaagen eingesetzt werden. Zwei ähnliche Prototypen, die ebenfalls auf Wägezellen basieren, wurden von Lessel et al. (2016) und Soubam et al. (2017) entwickelt. Da die Waage mit einem Smartphone direkt, energieeffizient kommunizieren sollte, war Bluetooth LE die beste Wahl. Eine Anforderung an den Prototyp war, dass neben dem Platzieren eines Objektes auf der Waage, keine weitere Interaktion mit der Waage notwendig sein sollte. Eine häufig umständliche Bluetooth-Kopplung sollte verhindert werden. Daher wurden Bluetooth-Advertisement-Pakete verwendet, die alle zwei

Sekunden ausgesandt wurden. Diese Art von Datenpaketen wird normalerweise für das Auffinden anderer Bluetooth-Geräte temporär versandt. Allerdings nutzen sie auch Apples iBeacon- oder Googles Eddystone-Protokoll für die Abschätzung der Distanz zwischen Sender und Empfänger. Für den Prototyp wurden eigene Pakete genutzt, bei denen das aktuell gemessene Gewicht als Payload versandt wurde. Die Waage konnte allerdings auch als Beacon eingesetzt werden, sodass eine App erkennen konnte, ob sich der Nutzer in deren Nähe befand, indem die RSSI-Werte für die Sende- und Empfangsstärke berücksichtigt wurden. Die Pakete wurden als Broadcast ausgesandt, sodass alle Geräte in der Nähe diese empfangen konnten. Somit konnte eine Waage mit mehreren Personen geteilt werden, wenn diese beispielsweise in der Wohnung stand. Da die Reichweite von Bluetooth LE beschränkt ist (ca. 10 m) und keine kritischen / persönlichen Daten per Funk versandt wurden, war keine Datenverschlüsselung notwendig, um die Privatsphäre zu wahren.

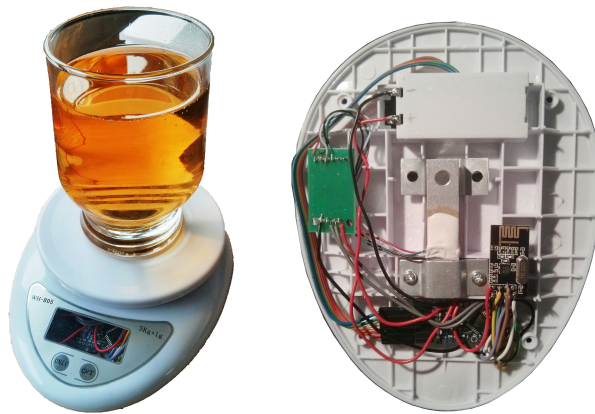


ABBILDUNG 8.14: Bilder vom zweiten Smartscale-Prototyp.

Für den **zweiten Prototyp** (siehe Abbildung 8.14) wurde eine kostengünstige Küchenwaage umgebaut, indem zunächst, bis auf die Wägezelle, die eingebaute Elektronik entfernt wurde. Daraufhin wurden ein Arduino Mini Pro, ein HX711-Analog-Digital-Wandler und ein nRF24L01+-Funkmodul in das Gehäuse verbaut. Das nRF24L01+-Funkmodul kann mit einer angepassten Software bluetoothkompatible Pakete aussenden. Der HX711-AD-Wandler wurde mit der eingebauten Wägezelle verbunden, die für ein Gewicht von bis zu 5 kg ausgelegt war. Die Stromversorgung wurde mit einem externen, schaltbaren Batteriegehäuse mit drei Mignon-Akkus bzw. Batterien gewährleistet.

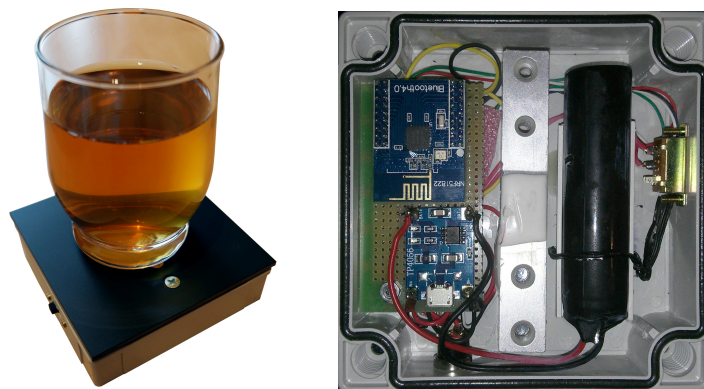


ABBILDUNG 8.15: Bilder vom dritten Smartscale-Prototyp.

Um die Mobilität im Hinblick auf die externe Stromversorgung und die Kompaktheit zu erhöhen, wurde ein **dritter Prototyp** (siehe Abbildung 8.15) entwickelt. Hierfür wurde die Größe reduziert ($10,0 \times 10,0 \times 3,2$ cm statt $16,2 \times 12,7 \times 3,3$ cm) und ein wiederaufladbarer 18650-Lithium-Akku mit 3,4 Ah inklusive Schutzschaltung und Ladeschaltung (TP4056) integriert. Der Akku

konnte mit einem Adapterkabel (Hohlstecker auf USB) per USB-Port oder Netzteil aufgeladen werden, wobei die Indikator-LEDs der Ladeschaltung von außen sichtbar waren. Für diesen Prototyp wurde Nordics NRF51822-Bluetooth-4.0-SoC eingesetzt. Dieser SoC ist im Vergleich zur Kombination aus Mikrocontroller und externem Funkmodul platzsparender und energieeffizienter. Der NRF51822-Controller bot jedoch nur beschränkte Arduino-Unterstützung, sodass die HX711-Bibliothek portiert werden musste, um den Code mit Nordics SDK verwenden zu können. Entsprechend der Anforderung „kostengünstig“, konnte der Preis für den zweiten Prototyp auf ca. 30 € und für den dritten auf ca. 40 € begrenzt werden. Die Preise waren vergleichbar mit denen von Lessel et al. (2016) und Soubam et al. (2017). Die Akkulaufzeit der eigenen Prototypen konnte nur begrenzt mit anderen Prototypen verglichen werden, da verschiedene Bauteile, Übertragungsmethoden sowie Energiesparmaßnahmen eingesetzt wurden. Der Prototyp von Lessel et al. (2016) lief ungefähr 31 Stunden, der von Soubam et al. (2017) je nach Einstellungen ca. 8 bzw. 21 Stunden. Die eigenen Prototypen erreichten eine durchgängige Laufzeit von ca. drei Wochen, hierbei wären jedoch noch weitere Verbesserungen durch Anpassung der Hard- und Software möglich gewesen.

Nutzung der Waage Die Waage diente dazu, das Gewicht einer Essensportion oder das Volumen (mittels der Dichte von Wasser) von Getränken zu bestimmen. Normalerweise werden Gefäße verwendet, um Lebensmittel auf der Waage zu platzieren. Daher muss das Leergewicht des jeweiligen Gefäßes abgezogen werden, um das Gewicht des Lebensmittels zu erhalten (Tara). In Kombination mit der entwickelten App, konnte der Nutzer die Reihenfolge festlegen, in der das Leergewicht bestimmt werden sollte. Zusätzlich war es möglich, das Gewicht eines bestimmten leeren Gefäßes zu speichern, um es direkt in der App auswählen zu können. Eine bestimmte Menge eines Gerichts oder Getränks konnte ebenfalls gespeichert werden. Hierbei wurden das Gewicht sowie die zugehörige Lebensmittelart als Favorit gesichert, was hilfreich sein konnte, falls bei einem bestimmten Gericht / Getränk die Menge ungefähr gleich bleibt. Dies ist beispielsweise der Fall, wenn Kaffee immer aus derselben Tasse getrunken wird. Dies erlaubt eine relativ präzise Mengenbestimmung, ohne die Waage wiederholt nutzen zu müssen.

Implementierung

Überblick Für den Prototyp des Nahrungserfassungssystems wurden die beiden Bluetooth-Waagen (Smartscale-Prototyp 2 und 3), ein Smartphone (Huawei Nexus 6P) mit Android 7.1 und eine runde Smartwatch (Motorola Moto 360 gen 1) mit „chin“ (Kante am unteren Rand), auf der Android Wear 1.5 ausgeführt wurde, verwendet.

Abbildung 8.16 links zeigt alle Geräte, die für den High-Fidelity-Prototyp verwendet wurden. Die Kommunikation zwischen den Geräten ist in Abbildung 8.16 rechts dargestellt.

Smartphones können durch ihre Größe eine wesentlich höhere Akkukapazität als Smartwatches besitzen (in diesem Fall 3.450 mAh vs. 320 mAh). Daher sollten energieintensive Aufgaben auf das Smartphone ausgelagert werden. Das Smartphone war aus diesem Grund für den Aufbau einer Internetverbindung per WiFi und das Empfangen der Bluetooth-Pakete der Waage zuständig. Internet wurde ausschließlich bei der Erstverwendung der optionalen Offline-Spracherkennung von Android benötigt, da hierfür ein Sprachmodell heruntergeladen werden musste. Androids Offline-Spracherkennung besaß eine Fehlerrate (word error rate) von 13,5 % (McGraw et al., 2016).

Um eine fließende Interaktion zwischen den Geräten zu ermöglichen, war eine Datensynchronisation zwingend notwendig und wurde bereits als Anforderung identifiziert. Die Synchronisation zwischen Smartscale, Smartphone und Smartwatch wurde per Bluetooth durchgeführt. Nach dem Empfang einer Update-Anfrage von der Smartwatch wurden vom Smartphone die Favoriten, der Verlauf (die neuesten 10 Einträge) und die Liste der Gefäße zur Smartwatch gesendet. Gewichtsdaten der Waage wurden vom Smartphone empfangen und anschließend direkt an die Smartwatch gesendet. Neue Einträge, die mit der Smartwatch angelegt wurden sowie Logs über die Nutzung der Smartwatch-App, wurden zum Smartphone weitergeleitet und in einer Datenbank gespeichert.



ABBILDUNG 8.16: Links: Geräte, die für den High-Fidelity-Prototyp verwendet wurden. Rechts: Kommunikation zwischen den Geräten.

Gemäß der identifizierten Anforderungen besaßen das Smartphone und die Smartwatch eine lokale Lebensmitteldatenbank, um die Auswahl der Art des Essens oder Getränks zu erleichtern. Die Datenbank wurde mit deutschen Lebensmittelnamen befüllt, die von verschiedenen Websites stammten. Die Datenbank der Smartwatch wurde auf die relevantesten Einträge manuell von 30.000 auf 7.000 Einträge reduziert, da es relativ mühsam war, auf der Smartwatch ohne Texteingaben und Spracherkennung eine große Anzahl an Einträgen zu durchsuchen. Erst Android Wear 2.0 erlaubte eine Texteingabe auf Android-Smartwatches per on-screen-Tastatur, die allerdings aufgrund ihrer kleinen Größe nur schwer verwendbar war. Die zu speichernden Daten wurden in einer zusätzlichen Datenbank auf dem Smartphone gesichert. Hierzu zählten: alle Gefäße, Lebensmitteleinträge (inklusive Bildern), Favoriten und Nutzungsdaten der Smartphone-/ Smartwatch-App. Die Datenbank-Tabelle mit den Gefäßen wurde beim ersten Start mit einigen Vorlagen befüllt, sodass diese direkt verwendet oder vom Nutzer editiert werden konnten. Eine Ergänzung von Gefäßen durch den Nutzer war ebenfalls möglich.

Nutzeroberfläche Abbildung 8.17 zeigt die wesentlichsten Bildschirmanzeigen (Screens) des Aufzeichnungsprozesses des Smartphones und der Smartwatch. Im Folgenden wird der Ablauf mit festen Nahrungsmitteln vorgestellt, der identisch mit Getränken ist.

Die Erfassung konnte bei beiden Geräten direkt über die Hauptbildschirme begonnen werden. Beim Smartphone war dies auf dem Home-Screen mittels Widget möglich, bei der Smartwatch wurde ein eigenes Ziffernblatt (watch face) gestaltet, das Abbildung 8.16 zeigt. Auf beiden Geräten konnte der Nutzer auswählen, ob ein Essen oder Getränk aufgezeichnet werden sollte. Zusätzlich war es möglich, das aktuell gemessene Gewicht der Waage anzuzeigen. Die Smartphone-App erlaubte zudem die Anzeige und das Editieren bereits gespeicherter Einträge, indem diese im Verlauf (history) berührt wurden (siehe Abbildung 8.17). Auf diese Weise war es auch möglich, unvollständige Einträge, die zuvor mit der Smartwatch erstellt wurden, zu vervollständigen. Entsprechende unvollständige Einträge wurden orange markiert. Generell mussten nur die Art und das Gewicht eines Lebensmittels vom Nutzer angegeben werden.

Auf beiden Plattformen war es möglich, mittels einer Swipe-Geste nach links oder rechts durch die verschiedenen Tabs zu wechseln. Auf dem Smartphone wurden zusätzlich Buttons zur Navigation angeboten. Das Abbrechen der Erfassung war möglich, indem der Zurück-Button von Android oder der Button „cancel“ (Abbruch) gedrückt wurde. Diese Buttons waren aus Platzgründen beim Smartwatch-Interface nicht enthalten. Es konnte dort allerdings der Hardware-Button gedrückt oder wiederholt die Swipe-Geste nach rechts genutzt werden, bis das Ziffernblatt wieder sichtbar war.

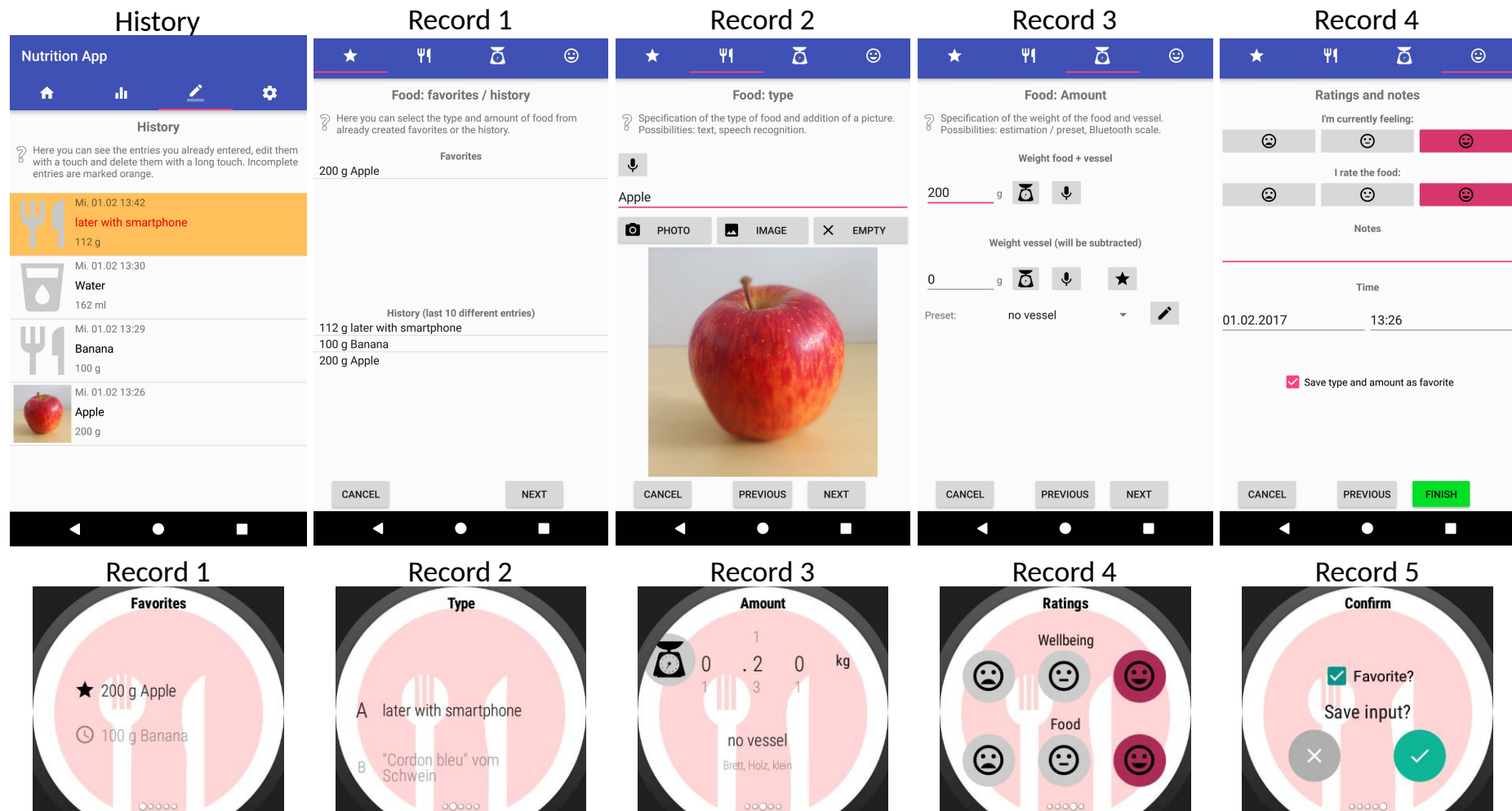


ABBILDUNG 8.17: Ausgewählte Bildschirmanzeigen des Smartphones (oben) und der Smartwatch (unten).

Nach dem Starten des Modus zum Erfassen von Gerichten wurde ein Bildschirm angezeigt, bei dem der Nutzer einen Eintrag aus eigenen Favoriten oder den zuletzt gespeicherten Gerichten auswählen konnte (siehe Abbildung 8.17 Record 1). Dies war bei der Smartwatch identisch gestaltet, allerdings waren Favoriten in der Liste mit jeweils einem Stern und Einträge aus dem Verlauf mit einem Uhren-Symbol markiert. Da auf dem Bildschirm von Smartwatches nur wenig Platz zur Verfügung steht, wurde der Modus zur Erfassung (Gericht / Getränk) mit einem entsprechenden Hintergrundbild verdeutlicht. Ein Eintrag konnte ausgewählt werden, indem dieser berührt wurde. Hierdurch wurde die Art des Essens, die Menge und die Art des Gefäßes bzw. dessen Gewicht automatisch ausgefüllt. Auf beiden Geräten wurde der letzte Bildschirm (Record 4 bzw. Record 5), nach Auswahl eines Favoriten bzw. Eintrags aus dem Verlauf, angezeigt.

Falls der Nutzer sich dazu entschied, keine Vorgabe zu nutzen, konnte zum nächsten Bildschirm (Record 2) gewechselt werden, wo die Art des Essens ausgewählt wurde. Auf dem Smartphone war es möglich, eine Texteingabe mit Autovervollständigung (eine Liste mit evtl. passenden Einträgen aus der Lebensmitteldatenbank wurde angezeigt) oder Spracherkennung unter Verwendung der Google-API, einzusetzen. Zusätzlich konnte ein Foto vom Nutzer aufgenommen oder ein bereits existierendes Bild aus dem Speicher ausgewählt werden. Die Smartwatch war bei ihren Eingabemethoden relativ stark eingeschränkt, daher konnten die Einträge hier ausschließlich aus einer Liste ausgewählt werden. Um diese Auswahl zu beschleunigen, wurden Einträge, die wahrscheinlich selten genutzt würden, aus der Lebensmitteldatenbank der Smartwatch entfernt. Zusätzlich konnte, wie bei einem Register, der erste Buchstabe des Lebensmittels in einer zusätzlichen Liste auf der linken Seite ausgewählt werden, um das Scrollen in der Liste zu beschleunigen. Falls der Nutzer ein bestimmtes Lebensmittel in der Liste nicht finden konnte oder in Eile war, war es möglich „later on smartphone“ (später auf dem Smartphone) auszuwählen. Dieser Eintrag war beim Erreichen des Bildschirms bereits vorausgewählt. Auf dem Smartphone konnte der Nutzer die Lebensmittelart später ergänzen, was hier wesentlich angenehmer war. Anschließend konnte zum nächsten Bildschirm (Record 3) gewechselt werden, auf dem die Menge bzw. das Gewicht eingegeben werden konnte.

Auf beiden Geräten war es möglich, das Gewicht des Essens inklusive Gefäß anzugeben. Das Gewicht konnte optional mittels der Smartscale gewogen werden. In diesem Fall musste der Nutzer das Essen nur auf der Waage platzieren und den entsprechenden Button auf dem Interface der Smartwatch oder dem Smartphone betätigen. Auf dem Smartphone standen zusätzlich die Spracherkennung und on-screen-Tastatur zur Verfügung. Der Gewichtswert konnte auf der Smartwatch manuell angepasst werden, indem drei Listen mit je einer Ziffer entsprechend eingestellt wurden. Das Gewicht des Gefäßes wurde vom Gesamtgewicht subtrahiert. Das Gefäßgewicht konnte festgelegt werden, indem es aus einer Liste mit Vorgaben und benutzerdefinierten Einträgen ausgewählt wurde. Auf dem Smartphone konnte das Leergewicht des Gefäßes direkt per Texteingabe, Spracherkennung oder mittels der Smartscale erfasst werden. Der Nutzer konnte hier zudem einen neuen Eintrag für ein Gefäß mit dem aktuellen Gewicht anlegen (Button mit Stern-Symbol). Ein Editor zum Anpassen der Gefäßliste war ebenfalls vorhanden. Auf der Smartwatch konnte mittels einer Liste ausschließlich aus bereits hinzugefügten Gefäßen ausgewählt werden. Falls der Anwender das Gewicht des Lebensmittels ohne Gefäß schätzen wollte oder ein Nahrungsmittel ohne Gefäß gewogen wurde (z. B. eine Frucht), konnten die Standardeinstellungen unverändert bleiben („no vessel“).

Auf der nachfolgenden Bildschirmanzeige konnten das allgemeine, aktuelle Wohlbefinden und das Gericht mit einer dreistufigen Skala (Smileys) bewertet werden. Auf dem Smartphone war es möglich, eigene Notizen zu ergänzen. Hier konnten zudem das Datum und die Uhrzeit verändert werden, falls eine nachträgliche Erfassung stattgefunden hat. Um auf der Smartwatch den Vorgang zu beschleunigen, wurde stets die aktuelle Systemzeit verwendet. Beim Smartwatch-Interface existierte eine zusätzliche Bildschirmanzeige (Record 5), um zu verhindern, dass Nutzer vertikal scrollen mussten, da dies nicht direkt ersichtlich war. Diese Anzeige erlaubte es, den Vorgang zu bestätigen oder abubrechen und verfügte über eine Checkbox, um den Eintrag als Favorit zu

speichern.

Nach dem Drücken des Buttons mit der Beschriftung „Finish“ bzw. mit einem Haken-Symbol (Smartwatch), war der Vorgang abgeschlossen und der Eintrag wurde gespeichert. Falls es bei der Smartphone-App fehlende, benötigte Felder gab (Lebensmittelart und Menge) oder das Gewicht des Gefäßes größer als das Gesamtgewicht inklusive Gefäß war, konnte der Nutzer den Vorgang nicht abschließen und es wurde mit einer Meldung auf dieses Problem hingewiesen.

Evaluation

Aufbau der Feldstudie Der High-Fidelity-Prototyp wurde mit Probanden getestet, die Interesse an der Nahrungserfassung zeigten und den Prototyp einen Tag lang ausführlich verwenden wollten. Zudem sollten die Probanden bereits ein Android-Smartphone mit Android 6 oder höher nutzen, da sie hierdurch bereits Erfahrung mit Android besaßen und zudem ihr Gerät für die App aktuell genug war. Ältere Android Versionen / Geräte hätten grundsätzlich ebenfalls funktionieren können, jedoch war Bluetooth 4.0 eine notwendige Anforderung zur Kommunikation mit der Smartscale. Die Bereitstellung eines entsprechenden Smartphones wäre zwar möglich gewesen, für diese Studie war es allerdings von Bedeutung, dass die Nutzer ihr eigenes, gewohntes Smartphone nutzen konnten und nicht ein zusätzliches oder fremdes Smartphone verwenden mussten. Hierdurch hätte die Aussagekraft der Ergebnisse stark gelitten.

Zunächst wurde den Probanden der vorgesehene Ablauf der Studie vorgestellt und erklärt, welche Daten von den verschiedenen Apps aufgezeichnet werden sollten. Danach wurde der Prototyp präsentiert und die Funktionen aller zur Verwendung stehenden Geräte erklärt. Im Weiteren wurden die Smartwatch und das Smartphone des jeweiligen Probanden miteinander per Bluetooth verbunden. Die Smartwatch musste hierbei in den Werkzustand versetzt und sodann die App neu installiert werden. Auf dem Smartphone wurde ebenfalls die zugehörige App installiert. Das spezielle Smartphone-Widget und das Ziffernblatt wurden daraufhin eingerichtet. Abschließend wurde eine Einführung in die Nutzung des Systems zusammen mit den Studienteilnehmern durchgeführt. Für alle Teilnehmer war die Smartwatch ein neues Gerät, das sie bisher weder getragen noch verwendet hatten.

Die Studienteilnehmer wurden gebeten, alle von ihnen verzehrten Nahrungsmittel / Getränke mittels ihrem Smartphone, der Smartwatch und Smartscale zu erfassen. Sie konnten die Nutzung bestimmter Geräte des Prototyps frei entscheiden. Am Ende der Studie wurde die Datenbank mit den gespeicherten Daten, im Beisein des Teilnehmers, vom Smartphone heruntergeladen. Zusätzlich wurden die Probanden in einem semistrukturierten Interview befragt. Hierbei wurden einige Fragen verwendet, die bereits beim Sammeln der Anforderungen genutzt wurden. Die Interviews dauerten ca. 15 Minuten pro Person und deckten 25 Fragen zu folgenden Bereichen ab: demografische Daten, Smartphone-Nutzung, Erfahrung mit Gesundheits-Apps, Bereitschaft Smartwatches zu nutzen, Bedenken gegenüber der Privatsphäre, Wissen über Hilfsmittel zur Nahrungserfassung, Motivation Lebensmittel zu erfassen und bevorzugter Ablauf bei der Erfassung.

Semistrukturierte Interviews Semistrukturierte Interviews wurden mit fünf Studienteilnehmern (Frauen: 1, Männer: 4) im Alter zwischen 24 und 31 (Mittelwert: 26,2) durchgeführt. Der Großteil der Probanden waren Studierende oder arbeiteten an der Universität. Zunächst wurden demografische Daten sowie Informationen über die Nutzung von mobilen Geräten und Nahrungserfassung erhoben.

Bezüglich der Nutzung von Smartphone-Apps berichtete eine Probandin (P5), dass sie bereits ihre Flüssigkeitsaufnahme mit der App („Plant Nanny“) erfasst hatte, um mehr zu trinken. Bilder von ihrem Essen nahmen die meisten Teilnehmer nur selten mit ihrem Smartphone auf. Vor allem ungewöhnlich „ästhetische“ Gerichte wurden fotografiert und beispielsweise Freunden oder Familienmitgliedern gezeigt. Hinsichtlich der Präzision bei der Mengenabschätzung war nur eine Person (P1) der Meinung, dass ihr dies bei Essen und Trinken relativ exakt gelingt. Die anderen

Studienteilnehmer gingen davon aus, dass sie die Menge fester und flüssiger Lebensmittel nur unpräzise abschätzen können.

Nach dem allgemeinen Teil des Interviews wurden die Teilnehmer gebeten, Kommentare über den Prototyp abzugeben. Hinsichtlich der Mengenabschätzung wurde die Waage als absolut notwendig angesehen, um mit einer hinreichenden Genauigkeit die Menge von Essen zu bestimmen. P2 bemerkte, dass die Waage „brillant“ sei, da sie die Eingabe der Menge stark vereinfachte. Er war zudem überrascht, wie ungenau er das Gewicht von Essen abschätzen konnte. Zum Beispiel ging er nicht davon aus, dass eine Orange (235 g) deutlich mehr wiegt als ein von ihm zubereitetes Sandwich (135 g).

Alle Teilnehmer besaßen im Hinblick auf die Wahrung der Privatsphäre dieselbe Meinung bezüglich der Nutzung der Smartwatch, Smartscale und Spracherkennung. P5 war besorgt über den Einsatz der Smartscale in der Öffentlichkeit: „ich nehme das Ding doch nicht zum Mc Donalds und wiege meinen Burger“. Die Smartwatch wurde von den Teilnehmern geschätzt, da diese anderen Personen die Nahrungserfassung nicht offenbart. P3 empfand es peinlich, wenn andere herausfinden würden, dass er die Nahrung erfasst. Die Spracherkennung wurde insbesondere von P2 geschätzt und von ihm zu Hause verwendet. Er würde sie allerdings nicht in der Arbeit verwenden: „... in der Arbeit, wenn ich gerade neben einem Kollegen stehe oder so, wäre es mir peinlich.“. Die Spracherkennung arbeitete bei den Wörtern, die er bei der Erfassung verwendete, korrekt. Andere Teilnehmer empfanden ebenfalls, dass die Spracherkennung in privaten Situationen eine hilfreiche Ergänzung darstellt. Dennoch wurde sie während der Studie nur selten eingesetzt.

Der Kontext beeinflusste die Gerätenutzung der Probanden. P4 bemerkte, dass die Smartwatch hilfreich sei, wenn man in Eile ist, da Details im Nachhinein auf dem Smartphone ergänzt werden konnten. Nur P1 empfand die Waage als weniger nützlich, da er die meiste Zeit während der Studie unterwegs war. Die anderen Teilnehmer platzierten die Waage in ihrer Wohnung. Sie konnten sich auch vorstellen, die Waage am Arbeitsplatz zu verwenden, allerdings nicht unterwegs.

Der Multi-Geräte-Aufbau zeigte Vorteile bei der Usability. Die Teilnehmer empfanden die Smartwatch (in Kombination mit der Waage) als besonders hilfreich, wenn wiederholt Gerichte eingetragen werden mussten, da in diesem Fall die Favoriten / der Verlauf verwendet werden konnten. Zusätzlich war die Smartwatch in Situationen nützlich, wenn die Probanden etwas in den Händen hielten bzw. das Smartphone nicht griffbereit war. Während die Smartwatch insbesondere bei wiederholten Eingaben Verwendung fand, wurde das Smartphone genutzt, um neue Lebensmittelarten und zusätzliche Details einzugeben. Die Smartphone-App wurde als nützlich eingestuft, da sie alle wesentlichen Bestandteile für die Nahrungserfassung beinhaltete.

Die Teilnehmer berichteten über keine Probleme bezüglich des Wechsels zwischen den Geräten. P2 merkte positiv an, dass die GUI der Smartwatch eine ähnliche Struktur wie die des Smartphones besaß. Die Bedeutung der GUI-Elemente war für ihn daher sofort verständlich.

Gemäß der aufgezeichneten Protokolle hatte P1 die Waage nicht zusammen mit der Smartwatch verwendet. P2 und P4 verwendeten die Smartscale öfters mit der Smartwatch und dem Smartphone. P3 erwähnte, dass er die Waage in Kombination mit der Smartwatch verwenden wollte, allerdings erst später den Button bemerkt hatte, um die Daten von der Smartscale auf die Smartwatch zu übertragen. Die anderen Teilnehmer äußerten hierzu keine Probleme.

Erinnerungen an die Nahrungsmittelerfassung durch die Smartwatch und das Smartphone wurden von P1 und P5 gewünscht. P5 hätte es begrüßt, wenn das System sie alle zwei bis drei Stunden regelmäßig an das Trinken erinnert hätte. Die anderen Teilnehmer erwähnten keine Probleme bezüglich Erinnerungen zur Erfassung von Essen und Trinken, sie wollten vielmehr weniger Erinnerungen, da sie störend wirken könnten. Sie meinten, dass sie nur Erinnerungen an Arbeitstagen (P2) oder passend zum Kontext (P3) vorteilhaft finden würden.

Das Ernährungserfassungssystem half den Teilnehmern dabei, über ihre Nahrungsaufnahme zu reflektieren. P2 meinte bezüglich der Nahrungsmittelbewertungen, dass die Bewertungsabfrage ihn auch darüber nachdenken ließ, ob es ihm gut oder schlecht geschmeckt hatte. Zusätzlich sah er sich später gerne die Fotos der Lebensmittel noch einmal an, die er bei der Erfassung aufgenommen

hatte. Da er bereits Nierensteine hatte, konnte er sich eine längerfristige Nutzung des Systems vorstellen, da das System ihn unterstützen könnte, ausreichend Flüssigkeit aufzunehmen.

Vorschläge für Erweiterungen

Ein Problem, das den Teilnehmern auffiel, war die hohe Anzahl an Lebensmittel-Datenbankeinträgen. Die Auswahl von Lebensmitteln war daher teilweise sehr zeitaufwändig und auf der Smartwatch kaum möglich, obwohl die Anzahl dort bereits reduziert wurde. P5 erwähnte, dass die Lebensmitteleinträge teilweise inkonsistent und verwirrend waren, da beispielsweise zum Teil Markennamen zusätzlich zum Produktnamen beinhaltet waren. Sie erwähnte außerdem, dass die Autovervollständigung viele andere Lebensmittel anzeigte, obwohl sie nur das Wort „Schokolade“ eintragen wollte, da es in vielen anderen Einträgen ebenfalls enthalten war. Auch P2 empfahl eine verbesserte Auswahlmöglichkeit der Lebensmittel. P1 und P4 wünschten sich eine Möglichkeit, Gerichte aus mehreren Zutaten ergänzen zu können.

Eine natürliche Spracheingabe wurde von P1 vorgeschlagen, da er die Zutaten eines Gerichts dem System während dem Kochen mitteilen wollte und hierbei die Hände nicht frei hätte, um mit den GUI-Interfaces zu interagieren. Diese Funktionalität war zu diesem Zeitpunkt auch in kommerziellen Apps nicht auffindbar (Franco et al., 2016), wäre allerdings eine sinnvolle Ergänzung gewesen. Mittlerweile existieren kommerzielle „Skills“ für Amazons Alexa, die in diese Richtung gehen, wie z. B. MyFitnessPal⁶ oder Track by Nutritionix⁷.

Diskussion

Die Studie zeigte, dass die Teilnehmer sich auch eine längerfristige Lebensmittelerfassung unter Verwendung mehrerer Geräte gut vorstellen konnten. Bei der Nutzung des Prototyps traten bei den Studienteilnehmern keine Probleme beim Wechsel zwischen Smartphone und Smartwatch auf. Die Smartwatch würde als zusätzliches Gerät in alltäglichen Situationen akzeptiert werden, während die Nutzung der Smartscale sich weitestgehend auf private Situationen beschränken würde. Spracherkennung führte zu vergleichbaren Problemen wie bei der Smartscale, da hierdurch ungewollte Aufmerksamkeit auf sich gezogen würde. Dennoch war die Smartscale in diesem System von Bedeutung, da auch einige Studienteilnehmer das Gewicht von Nahrungsmitteln nicht hinreichend genau abschätzen konnten. Sollte eine Nahrungsmittelerfassung diskret durchgeführt werden, war die Smartwatch hierfür am besten geeignet.

Erinnerungen bezüglich der Nahrungsaufnahme könnten kombiniert zu den Erinnerungen an die Erfassung gegeben werden. Möglicherweise besaß das angepasste Ziffernblatt der Smartwatch bereits eine ambiente Erinnerungsfunktion, da in der Arbeit von Y. Kim et al. (2016) Sticker ausreichend waren, um die Studienteilnehmer unbewusst zu erinnern.

Die Interviews deuteten darauf hin, dass Fotos insbesondere von speziellen, selbst gemachten Gerichten erstellt wurden. Es zeigte sich zudem, dass das Erfassen selbst gekochter Gerichte mit dem Prototyp schwieriger war, da bei Fertiggerichten in der Regel alle benötigten Daten auf der Verpackung abgedruckt sind. Für eine detaillierte Erfassung aller Einzelzutaten bei eigenen Gerichten wäre eine Erweiterung der Apps notwendig. Die Smartscale könnte beim Kochen als Ersatz für eine reguläre Küchenwaage eingesetzt werden und die Protokollierung der Zutatenmengen erleichtern. Eine zusätzliche Kombination mit einem Sprachassistenten wäre u. a. für eine vollständig berührungslose Systeminteraktion beim Kochen denkbar.

Zusammenfassung

In diesem Abschnitt wurde ein System für eine persönliche Nahrungsmittelerfassung vorgestellt, bei dem die jeweiligen Stärken mehrerer mobiler Geräte ausgenutzt werden konnten, um die

⁶<https://blog.myfitnesspal.com/say-hello-to-myfitnesspal-on-alexa/> (24.10.2019)

⁷<https://www.nutritionix.com/blog/2016/06/06/track-by-nutritionix-amazon-alexa-skill/> (24.10.2019)

Erfassung zu erleichtern und eine genaue Mengenbestimmung zu ermöglichen. Das System kombinierte ein Smartphone, eine Smartwatch und eine Smartscale und erlaubte die Nutzung der Eingabemodalitäten Berührung und Sprache. Ebenfalls waren tangile Eingaben durch das Platzieren von Lebensmitteln auf der Waage möglich. Auf diese Weise erlaubte der Prototyp eine größere Flexibilität als andere Systeme, die normalerweise ausschließlich auf ein Smartphone als Interaktionsgerät setzen. Eine Feldstudie zeigte, dass die Probanden alle drei Geräte nutzten und deren spezifische Vorteile situationsabhängig einsetzen konnten.

Künftig könnten längere Studien mit dem System zusätzliche Informationen über die Nutzung in verschiedenen Situationen erbringen. Weiterhin wären Erweiterungen und Verbesserungen des Systems möglich. Hierzu gehören intelligente Erinnerungen, die verschiedene Ausgabemodalitäten nutzen könnten. Wie bereits Franco et al. (2016) in ihrer Arbeit erwähnten, wären auch die Integration von natürlicher Sprachverarbeitung und Bilderkennung hilfreiche Ergänzungen. Eine weiterführende Studie bei der mehrere Eingabemethoden verglichen werden, wäre ebenfalls sinnvoll. Ein Vergleich von Bilderkennung und Texteingaben wurde beispielsweise bereits in der Arbeit von Lim et al. (2017) durchgeführt und könnte daher als Grundlage dienen.

8.3.5 Erkennung von Trinkvorgängen mittels einer Smartwatch und interaktivem / kooperativem Lernen

Einleitung

Für viele gesundheitsbezogene Anwendungen ist die Erkennung von Nutzeraktivitäten eine wesentliche Informationsquelle. Das Ziel des in diesem Abschnitt beschriebenen und in (Flutura, Seiderer, Aslan, Dang, Schwarz, Schiller & André, 2018) publizierten Projekts war es, zu erkennen, wann eine Person trinkt. Dies sollte mittels der Beschleunigungsdaten einer Smartwatch erkannt werden. Das System wurde als „DrinkWatch“ bezeichnet. Die Besonderheit hierbei war, dass die Erkennung in Echtzeit auf der Smartwatch durchgeführt und das verwendete Klassifizierungsmodell während der Nutzung personalisiert wurde. Die Personalisierung nutzte hierbei die Prinzipien des interaktiven und kooperativen maschinellen Lernens (engl. interactive / cooperative machine learning). Hierfür wurde der Nutzer in den Anpassungsvorgang einbezogen, indem dieser vom initialen Grundmodell erkannte mögliche Trinkbewegungen bestätigte oder als fehlerhaft rückmeldete. Dies ermöglichte eine Optimierung des Grundmodells und sorgte zu einem für den Nutzer transparenten Lernvorgang. Während der Umsetzung dieser beispielhaften Anwendung und deren Evaluation wurden die Möglichkeiten und Einschränkungen der mobilen Nutzung von interaktivem und kooperativem maschinellen Lernen deutlich erkennbar und werden nachfolgend angesprochen.

Verwandte Arbeiten

Aktivitätserkennung Zu den klassischen Aktivitäten, die bei der menschlichen Aktivitätserkennung (engl. human activity recognition = HAR) erkannt werden, zählen das Laufen, Rennen, Joggen, Stehen, Sitzen, Treppensteigen (auf / ab) und mit dem Lift fahren. Hierfür können Smartphones eingesetzt werden, da sie meist in der Hosentasche getragen werden und die Bewegung des Torso ausreichend ist, um typische Muster in den Daten des 3-Achsen-Beschleunigungssensors (Accelerometer) zu erfassen. Aktivitäten, die jedoch primär mit den Armen oder Händen durchgeführt werden, sind für die Smartphone-Sensorik nicht direkt erfassbar. Hierfür eignen sich allerdings Fitness-Tracker oder Smartwatches. Diese verfügen ebenfalls über Beschleunigungssensoren und werden am Handgelenk getragen. Damit ermöglichen sie beispielsweise, die typischen Bewegungen beim Essen, Trinken oder auch Rauchen zu erkennen (Shoaib et al., 2016). Sie können ausgewertet und für eine Beurteilung von gesundheitsbezogenen Verhaltensweisen genutzt werden. Eine direkte Erkennungsmöglichkeit von Aktivitäten mittels Smartwatch-Hardware wurde bereits in den Arbeiten von Bhattacharya und Lane (2016) und Shahmohammadi et al. (2017)

umgesetzt. In der eigenen Arbeit wurde zusätzlich interaktives und online maschinelles Lernen implementiert und vollständig auf einer Smartwatch ausgeführt.

Interaktives und kooperatives maschinelles Lernen Interaktives und kooperatives maschinelles Lernen gewinnt in der Aktivitätserkennung zunehmend an Bedeutung und wurde auf mobilen Geräten beispielsweise von Miu et al. (2015) sowie Shahmohammadi et al. (2017) eingesetzt.

Interaktives maschinelles Lernen (iML) unterscheidet sich vom klassischen, automatischen maschinellen Lernen dahingehend, dass der Nutzer direkt in den Prozess einbezogen wird („human-in-the-loop“) (Holzinger, 2016). Laut Holzinger (2016) besitzt iML seine historischen Ursprünge in „Reinforcement Learning“ (1950), „Preference Learning“ (1987) und „Active Learning“ (1996). Das Trainieren des Modells ist bei iML Teil des Systems und wird nicht ausschließlich während der Entwicklung durchgeführt. Daher besitzt iML großes Potenzial bei Anwendungen, in denen sich das System kontinuierlich an den Nutzer anpassen muss und teilweise das Feedback des Nutzers für Datenbeschriftungen benötigt wird. Ware et al. (2001) beschrieben ein System mit maschinellen Lernen bei dem vom Anwender eine gewisse Expertise für die Interaktion mit dem System erwartet wird. Dies wurde beim eigenen entwickelten System vom Nutzer nicht gefordert. Beim System von Fails und Olsen Jr (2003) war Expertise in maschinellen Lernen ebenfalls keine Notwendigkeit, da das interaktive maschinelle Lernen mit Designern evaluiert wurde. Bei der eigenentwickelten DrinkWatch wurde allerdings als Anwendungsbeispiel ein anderes Szenario gewählt und das System in Echtzeit auf einer Smartwatch ausgeführt.

In der Arbeit von Shahmohammadi et al. (2017) wurde interaktives maschinelles Lernen basierend auf Smartwatch-Sensordaten für Aktivitätserkennung evaluiert und hierbei festgestellt, dass bereits eine geringe Anzahl an Trainings-Daten ausreichend ist, um hohe Erkennungsraten zu erreichen. Die mit dem Verfahren personalisierten Modelle zeigten zudem wesentlich bessere Ergebnisse als die statischen Modelle. In der eigenen Arbeit wurden im Gegensatz zu (Shahmohammadi et al., 2017) „online learning“-Verfahren eingesetzt, bei denen während der Laufzeit das Modell inkrementell adaptiert werden konnte. Daher war es in diesem Fall nicht nötig, Daten an einen leistungsfähigeren Rechner zum Trainieren weiterzuleiten. Es wurden bei den Tests in der eigenen Arbeit zudem ausschließlich Daten verwendet, die während natürlichem Trinkverhalten (ohne Instruktionen) aufgezeichnet wurden.

Ähnlich wie bei interaktivem maschinellen Lernen wird bei **kooperativem maschinellen Lernen** (engl. cooperative machine learning = cML) versucht, ein Problem des maschinellen Lernens durch eine Zusammenarbeit zwischen Nutzer und System zu lösen. Bei cML liegt jedoch der Fokus auf der Reduktion des Aufwands, der bei der manuellen Annotation (Datenbeschriftung) von aufgezeichneten Daten auftritt (Z. Zhang et al., 2014). Bei cML ist eine enge Verbindung zwischen Annotationssoftware und einem Framework für maschinelles Lernen notwendig. In diesem Fall wurde die Software „NOVA“ von Baur et al. (2015) eingesetzt, die für maschinelles Lernen das SSI-Framework (Wagner et al., 2013) nutzt. Das Annotationsprogramm wird zunächst klassisch verwendet, sobald jedoch einige Daten markiert und mit einer Beschriftung versehen wurden, kann mittels SSI ein Modell erstellt und auf den Rest der Daten angewandt werden. Diese Vorschläge können anschließend vom Nutzer überprüft und gegebenenfalls mit weiteren Daten ein neues Modell erzeugt werden. Der Aufwand bei der Annotation wird hierdurch in vielen Fällen deutlich reduziert.

Active Learning Laut Holzinger (2016) besteht die Grundidee bei „active learning“ darin, dass ein maschineller Lernalgorithmus mit weniger Datenbeschriftungen als beim klassischen maschinellen Lernen eine höhere Genauigkeit erreichen kann, wenn dieser die Daten für den Lernprozess selbst auswählen kann. Ein solcher Algorithmus fragt normalerweise ein „Orakel“, welches ihm die korrekte Beschriftung der Daten mitteilen kann. Dies kann, wie auch bei der DrinkWatch, durch eine Person erfolgen.

Miu et al. (2015) präsentierten ein echtzeitfähiges Active-learning-Framework und untersuchten, wie Annotationen vom Nutzer gesammelt werden können, um ein personalisiertes HAR-Modell zu erstellen. Es wurde hierbei gezeigt, dass HAR-Modelle während der Verwendung erstellt werden können und keine Überwachung durch Experten oder nachträgliche Datenannotation benötigt wird. Von Miu et al. (2015) wurde das GUI des Smartphones verwendet, bei der DrinkWatch das Interface der Smartwatch.

Active learning wurde für verschiedene Arten der Klassifikation, mit verschiedenen Verfahren (z. B. SVMs (Tong & Koller, 2001)) und Nutzerabfragearten eingesetzt. Ein Überblick wird in den technischen Reports von Settles und seinem Buch (Settles, 2012) gegeben. Die am häufigsten verwendete Abfragemethode ist „uncertainty sampling“ (Lewis & Gale, 1994) oder „query on uncertainty“. Hierbei werden ausschließlich Datenbeschriftungen vom Nutzer abgefragt, bei denen der Klassifikator nur eine geringe Erkennungssicherheit besaß. Dieses Verfahren wurde bei der DrinkWatch eingesetzt. Gemäß Lewis und Gale (1994) zeigte dieses Verfahren ein besseres Ergebnis als „relevance sampling“, bei welchem dem Nutzer nur Daten zum Annotieren gegeben werden, die vom Klassifizierer bereits mit hoher Wahrscheinlichkeit korrekt zugeordnet wurden. Eine weitere Methode ist „query by committee“ (Seung et al., 1992), wobei mehrere deutlich unterschiedliche Modelle eingesetzt werden, die ein Komitee (engl. committee) bilden und über die Klassen von Daten zusammen abstimmen. Die Daten, bei denen sich die einzelnen Modelle am unsichersten sind, werden dem Orakel zur Beschriftung gegeben. Es existieren noch weitere Abfragemethoden, die allerdings stärker auf eine Fehlerreduktion fokussiert sind (Settles, 2012).

Implementierung

Zum Zeitpunkt der Entwicklung der DrinkWatch waren nur sehr wenige Systeme mit Smartwatch in der Literatur aufzufinden, die iML (Shahmohammadi et al., 2017; Shoaib et al., 2016) oder Online-Lernen (Miu et al., 2015) nutzten. Im Gegensatz zur eigenen Arbeit kombinierte jedoch keines der Systeme beide Ansätze und setzte die Verfahren direkt auf der Smartwatch ein.

Für den DrinkWatch-Prototyp wurde eine Smartwatch ausgewählt, da diese die Armbewegungen an ihrer Position messen kann und immer in Reichweite ist, wenn eine Interaktion mit dem Nutzer benötigt wird. Die Darstellungsmöglichkeiten der kleinen Smartwatch-Displays sind allerdings eingeschränkt. In diesem Fall wurde jedoch keine komplexere grafische Oberfläche benötigt. Im Idealfall wären nach einer ersten Anpassungsphase an die nutzerspezifischen Trinkbewegungen nur noch wenige Interaktionen nötig, um die Trinkvorgänge zu erfassen.

Sobald die DrinkWatch potenziell geeignete Daten einer Trinkbewegung (Armbewegung) zum Lernen registrierte, wurde der Nutzer gefragt, ob er getrunken hat oder nicht (binäre Klassifizierung). Dies war insbesondere dann sinnvoll, wenn leicht verwechselbare Bewegungen, wie z. B. Naseputzen oder Mundabwischen vom Trinkvorgang, nicht unterschieden werden konnten. Der Nutzer konnte zudem die Anfrage ignorieren. Eine Antwort des Nutzers wurde als Annotation / Beschriftung für die erfassten Daten verwendet und beeinflusste somit das maschinelle Lernen.

Der DrinkWatch-Prototyp bestand aus drei Komponenten, dem GUI, um Annotationen zu erfassen und erfasste Aktivitäten anzuzeigen, dem Modul zur **Datenerfassung und Signalverarbeitung** sowie dem Modul für **interaktives maschinelles Lernen**.

Hardware Als Smartwatch wurde eine Asus ZenWatch 2 mit Android Wear 2 eingesetzt, die ein rechteckiges 1,63 Zoll Display besaß. Sie verfügte über einen SoC (Qualcomm Snapdragon 400 APQ8026) mit einer 1,2 GHz ARM-Cortex-A7-Vierkern-CPU, Bluetooth 4.1 und WiFi, 512 MiB Ram, 4 GiB eMMC Speicher, einen Akku mit 400 mAh sowie eine USB-Schnittstelle zum Laden und Datenaustausch. Als Sensoren waren ein 3-Achsen-Accelerometer und Gyroskop sowie ein Mikrofon vorhanden. Neben dem Display waren als Aktuatoren ein Lautsprecher und Vibrationsmotor integriert. Die Smartwatch war zudem vergleichsweise preisgünstig (ca. 110 €) und verfügte über alle notwendigen Funktionen für den DrinkWatch-Prototyp.

Für die Datenannotation wurden u. a. Gewichtsdaten des eigenen Smartscale-Prototyps (Prototyp 3) (siehe Abschnitt 8.3.4) mit der Smartwatch per Bluetooth empfangen und aufgezeichnet.

GUI Die grafische Nutzerschnittstelle war minimal gestaltet und erlaubte es daher, auch Nutzern ohne Expertenwissen das System zu verwenden. Abbildung 8.18 zeigt die GUI der DrinkWatch. In der ersten Zeile sind zwei Buttons zum Starten und Stoppen des Systems erkennbar. Darunter zwei für die Auswahl, ob es sich um eine Trinkaktivität handelte: „Ja“ (Yes) oder „Nein“ (No).

Sobald eine Bestätigung des Nutzers nach einer Armbewegung notwendig war, wurde eine textuelle Meldung angezeigt („Have you been drinking?“ statt „Waiting on Event“) und der Standard-Systemklang abgespielt. Die Vibrationsfunktion wurde deaktiviert, um Verfälschungen der Beschleunigungsdaten zu vermeiden.



ABBILDUNG 8.18: GUI der DrinkWatch. (Abbildung verändert aus Flutura et al. (2018))

Warmstartmodell Es wurde ein Klassifizierungsmodell erstellt, das eine grundlegende Erkennung der Klassen „Trinken“ und „nicht Trinken“ ermöglichte. Damit konnte das System unmittelbar genutzt werden, ohne einen aufwändigeren Trainingsvorgang mit Nutzerdaten zu benötigen. Dieses sogenannte „Warmstartmodell“ konnte durch inkrementelles Lernen an den Nutzer angepasst (fine-tuning) bzw. personalisiert werden. Für das Warmstartmodell musste zunächst ein Datenkorpus aufgezeichnet und annotiert werden. Es wurden Bewegungsdaten in 25 Sitzungen von einer (22 Sitzungen) bis drei Personen mit den Smartwatches in verschiedenen Alltagssituationen, wie z. B. bei der Arbeit am Computer oder während eines Gesprächs, aufgezeichnet. Die Teilnehmer verfügten jeweils über ein gefülltes Trinkgefäß, aus dem sie bei Bedarf tranken. Die ZenWatch-2-Smartwatches wurden an der dominanten Hand getragen, mit der auch das Gefäß gegriffen wurde. Für die Annotation wurden Videoaufnahmen angefertigt und in 20 Sitzungen zusätzlich die Daten des Smartscale-Prototyps aufgezeichnet. Insbesondere die Gewichtsdaten der Smartscale erleichterten die Datenannotation der Trinkbewegungen im Zeitbereich immens. Das Trinkgefäß wurde nur während des Trinkvorgangs von der Waage genommen, sodass die Smartscale während des Trinkens immer ein Gewicht von 0 g erfasste.

Die Beschleunigungsdaten der Smartwatch und Gewichtsdaten der Smartscale wurden auf den Smartwatches aufgezeichnet. Die Datenmenge umfasste insgesamt 16 Stunden und 30 Minuten. Die Abtastrate der Beschleunigungssensoren der Smartwatch betrug 25 Hz. Es wurden die vom Android-System bereitgestellten linearen Beschleunigungsdaten verwendet, bei denen die Schwerkraft bereits kompensiert worden war. Bei einer Fenstergröße von 8 Sekunden und einer Überlappung von 7 Sekunden waren im Datenkorpus 5117 Datenfenster für Trinkaktivitäten und 26.288 für Nicht-Trinkaktivitäten vorhanden. Diese Daten wurden für das kooperative maschinelle Lernen und das Warmstartmodell genutzt. Um die Datenmenge der beiden Klassen beim Trainieren anzugleichen, wurde „random under-sampling“ eingesetzt.

Maschinelles Lernen Das Hauptaugenmerk lag beim DrinkWatch-Prototyp auf der Komponente für das maschinelle Lernen und der Abfrage der hierfür benötigten Daten. Der Nutzer sollte

durch Abfragen nur gestört werden, wenn der Informationsgewinn für den Klassifikator hierdurch möglichst hoch war (Amershi et al., 2014).

Der DrinkWatch-Prototyp verwendete für die Signalverarbeitung und das maschinelle Lernen mobileSSI (siehe Abschnitt 4.1.2). MobileSSI verfügte bereits über klassische Verfahren des maschinellen Lernens und wurde für diesen Prototyp um Online-Lernen erweitert.

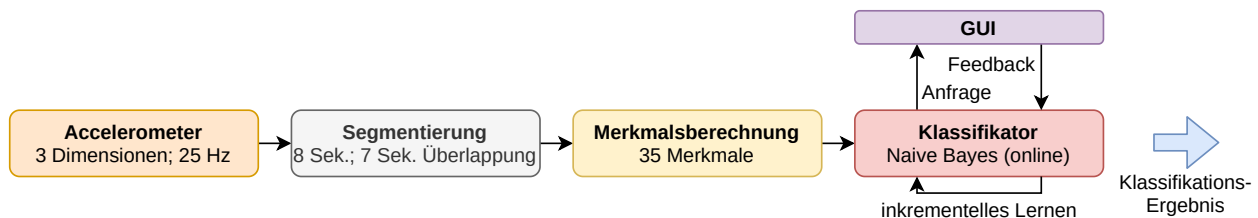


ABBILDUNG 8.19: Übersicht über die Signalverarbeitung und das interaktive maschinelle Lernen des DrinkWatch-Systems.

Abbildung 8.19 zeigt eine Übersicht hinsichtlich Signalverarbeitung und interaktivem maschinellen Lernen des DrinkWatch-Systems. Für die Echtzeitsignalverarbeitung mussten zunächst die Daten vom Beschleunigungssensor (drei Dimensionen, 25 Hz) der Smartwatch erfasst werden. Hierfür wurde die Android-Sensor-API eingesetzt, die nur unter Java vollständig nutzbar war. Die Daten der Smartscale wurden per Bluetooth von der Smartwatch empfangen und im Java-Code vom Android-System abgerufen. Diese Daten wurden an das in C++ geschriebene mobileSSI, per JNI-Schnittstelle zur Synchronisation und weiteren Verarbeitung eingespeist. Hier fand zunächst eine **Segmentierung** in 8-Sekunden-Fenster mit einer Überlappung von 7 Sekunden statt. Diese Fenstergröße war in den meisten Fällen ausreichend, um den kompletten Bewegungsablauf beim Trinken zu erfassen und verursachte bei den Systemrückgaben an den Nutzer nur relativ geringe Verzögerungen.

Der eingesetzte Klassifikator wurde mit **Merkmalsvektoren** genutzt. In der HAR-Literatur finden sich hierfür diverse Merkmale: siehe beispielsweise Bao und Intille (2004), Huynh und Schiele (2005), Lara und Labrador (2012) und Ravi et al. (2005). Hiervon wurde für die DrinkWatch eine Auswahl implementiert. Auf jeder Achse der Beschleunigungsdaten wurden folgende Merkmale bestimmt: Mittelwert, Standardabweichung, Varianz, Energie, Interquartilsabstand (engl. interquartile range = IQR), mittlere absolute Abweichung (engl. mean absolute deviation), quadratischer Mittelwert (engl. root mean square = RMS), Minimum, Maximum. Über alle Achsen wurden zusätzlich die Korrelation (XY, XZ, YZ), Mittelwert, Standardabweichung, Minimum, Maximum und IQR über die Länge jedes Sample-Vektors über alle Achsen (Magnitude) ermittelt. Insgesamt wurden 35 Merkmale über die 8-Sekunden-Fenster berechnet. Die Merkmale wurden zwischen 0 und 1 normalisiert.

Jedes 8-Sekunden-Fenster wurde mittels eines **Klassifikators** in Echtzeit analysiert und dessen Klasse (trinken / nicht trinken) bestimmt. Um das hierfür eingesetzte Modell mit den Nutzerfeedbacks verbessern zu können, musste ein maschinelles Lernverfahren verwendet werden, das neue Daten möglichst schnell in das Modell einbeziehen kann. Dies musste auch auf der limitierten Hardware der Smartwatch möglich sein. Es wurde daher ein Naive-Bayes-Klassifikator eingesetzt, der problemlos für Online-Lernen angepasst werden kann (siehe z. B. MOA von Bifet et al. (2010)). Die Klassifikation resultierte hierbei in Konfidenzen, die zur Entscheidung genutzt werden konnten, ob Nutzerfeedback eingeholt werden sollte. Eine mögliche Alternative stellte eine für Online-Lernen angepasste Variante einer linearen SVM (Tsai et al., 2014) dar. Für die durchgeführten Untersuchungen wurde die lineare SVM von LibLinear (Fan et al., 2008) ohne Online-Lernen, eingesetzt.

Der Lernvorgang des Systems begann mit einem Warmstartmodell, das durch das inkrementelle Lernen verfeinert wurde. Das „active learning“ wurde mit der „query on uncertainty“ Abfrage

umgesetzt. Die maximale Akkulaufzeit dieses Systems bei kontinuierlicher Ausführung auf den ZenWatch-2-Smartwatches betrug ca. 4 Stunden.

Evaluation

Naive-Bayes im Vergleich zur linearen SVM Für die statischen Modelle wurden der Naive-Bayes- und lineare SVM-Klassifikator miteinander verglichen. Evaluiert wurde auf dem Datensatz mit einer festen Aufteilung zwischen Trainings- und Test-Daten. Diese Test-Daten wurden auch für die Simulation des inkrementellen Lernens verwendet. Die Ergebnisse sind in Tabelle 8.3 dargestellt. Es wurde bei Naive-Bayes eine Treffergenauigkeit von 76,5 %, bei der SVM 82,3 % erreicht. Die Erkennungsraten der Nicht-Trinkaktivitäten waren geringer als die von den Trinkaktivitäten. Dies ist bei binären Klassifikationen mit meist sehr heterogenen Restklassen zu erwarten.

	Naive-Bayes	lineare SVM
Trinken	81,4 %	84,9 %
¬ Trinken	71,6 %	79,8 %
Mittelwert	76,5 %	82,3 %

TABELLE 8.3: Treffergenauigkeiten bei Modellen, die mit einer festen Teilung zwischen Trainings- und Test-Daten trainiert bzw. evaluiert wurden.

Simulation des inkrementellen Lernens Um das inkrementelle Lernen zu untersuchen, wurde dieser Vorgang mit der Annotationssoftware NOVA simuliert, die kooperatives maschinelles Lernen enthielt. Es wurde zunächst ein Basismodell mit acht annotierten Datenbereichen mit Trinkaktivität trainiert. Dieses Modell wurde daraufhin eingesetzt, um die anderen Bereiche zu klassifizieren. Falls die mit den Klassifikatoren von NOVA bestimmten Konfidenzen innerhalb eines bestimmten Intervalls lagen, annotierte der Nutzer den Datenbereich manuell. Danach wurde das Modell weitertrainiert. Falls die Konfidenzen über dem Intervall lagen, wurden die Datenbereiche als korrekt erkannt angenommen. Dieser Vorgang wurde für die gesamten Trainings-Daten wiederholt. Die Modelle wurden nach jedem Trainingsvorgang, mit den für das Training nicht genutzten Test-Daten, evaluiert.

Die Simulation wurde mit zwei verschiedenen Konfidenz-Intervallen ($[0,5; 0,7]$ und $[0,7; 0,9]$) und zwei Klassifikatoren durchgeführt (lineare SVM von LibLinear und Naive-Bayes) (siehe Abbildung 8.20). Es zeigte sich, dass der Naive-Bayes-Klassifikator anfangs wesentlich instabiler war als die lineare SVM. Dies könnte damit zusammenhängen, dass der Naive-Bayes-Klassifikator zunächst empfindlicher auf verändertes, personenabhängiges Verhalten gegenüber dem Basismodell reagierte. Nach einer Ergänzung von ca. 30 Beschriftungen (Entscheidungen: „trinken“ / „nicht trinken“) stabilisierte sich das Verhalten allerdings. Gemäß dem Verlauf der Treffergenauigkeiten in der Abbildung schien bei Naive-Bayes das höhere Konfidenz-Intervall ($[0,7; 0,9]$) besser geeignet zu sein, da hierbei weniger große Schwankungen in der Treffergenauigkeit auftraten. Bei der linearen SVM war im Gegensatz hierzu eher das niedrigere ($[0,5; 0,7]$) Intervall vorteilhaft.

Interaktives maschinelles Lernen Die DrinkWatch wurde mit zwei Nutzern getestet, die bei der Interaktion mit dem System mündliches Feedback geben sollten. Die beiden Studienteilnehmer interagierten mit dem System über eine Stunde hinweg. Gemäß der vorherigen Simulation wurde der Naive-Bayes-Klassifikator mit dem höheren Konfidenz-Intervall ($[0,7; 0,9]$) eingesetzt, da dieses sich potenziell schneller stabilisierte. Für das (Warmstart) Basismodell wurden hierbei die Daten aus 40 annotierten Bereichen genutzt. Die Qualität der Modelle konnte dadurch bestimmt werden, wie häufig das System nachfragte und ob eine Nachfrage angemessen war.

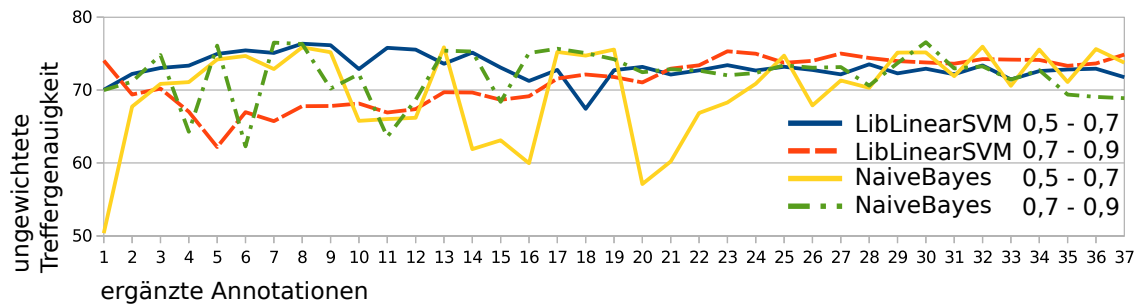


ABBILDUNG 8.20: Trainingsfortschritt mit verschiedenen Konfidenzen und Modellen.

Beide Nutzer waren der Meinung, dass die Trinkaktivitäten meistens akkurat erkannt wurden. Es wurden jedoch auch eine Vielzahl anderer Bewegungen als Trinkaktivität eingestuft, was die Nutzer als störend empfanden. Die relativ hohe Anzahl an falsch bewerteten anderen Bewegungen beruhten sehr wahrscheinlich auf dem großen Ungleichgewicht zwischen den beiden Klassen und war bereits in der Evaluation erkennbar, da die Erkennungsrate der Nicht-Trinkaktivitäten geringer war.

Das Systemverhalten war für die beiden Nutzer insgesamt gut nachvollziehbar. Sie bemerkten, dass eine Trinkbewegung meistens eindeutig erkannt wurde, wenn das Trinkgefäß langsam und stetig bewegt wurde. Diese Aktivitäten konnten von den Teilnehmern nachgeahmt sowie beschrieben werden und lösten beim System zuverlässig die Erkennung einer Trinkaktivität aus.

Zusammenfassung

In dieser Arbeit wurden Lösungen für die Probleme des klassischen maschinellen Lernens bezüglich der Erfassung authentischer Daten aus dem Alltag sowie einer fehlenden Transparenz und Nutzerkontrolle vorgestellt. Hierfür wurden mobiles interaktives und kooperatives maschinelles Lernen eingesetzt, wodurch Nutzer ihre eigenen Aktivitäten annotieren konnten und etwas Kontrolle über den Prozess des maschinellen Lernens erhielten. Kooperatives maschinelles Lernen stellt ein vergleichsweise junges Forschungsgebiet dar und diese Arbeit sollte Limitationen und das Potenzial für zukünftige Anwendungen aufzeigen.

Mit der DrinkWatch wurde ein mobiles System entwickelt, das mit der Nutzung von interaktivem und kooperativem maschinellen Lernen erlaubt, Trinkaktivitäten mit einer Smartwatch zu erfassen. Eine grundsätzliche Verbesserung des Klassifizierungsmodells war durch eine Nutzerantwort (Beschriftung) an das System möglich. Es erlaubte, die Protokollierung von Getränken zu erleichtern und dabei unter Nutzung von maschinellen Lernen die hierbei häufig gewünschte persönliche Privatsphäre zu schützen und Kontrolle über das System zu ermöglichen. Die Daten der Smartscale konnten als verlässliche Basis (Ground Truth) für die Datenannotationen eingesetzt werden. Somit erleichterte sie die Implementierung des Grundmodells (Warmstartmodell) maßgeblich.

Die in dieser Arbeit genutzten Verfahren wurden beispielhaft zu einer Erkennung von Trinkaktivitäten eingesetzt. Sie wären auch auf andere Probleme bei der Erkennung menschlicher Aktivitäten übertragbar. Es zeigte sich, dass moderne Smartwatches ausreichend Leistung besitzen, ein solches System nutzen zu können. Das System wäre auch direkt auf einem Android-Smartphone oder Tablet lauffähig gewesen. Die bei der DrinkWatch eingesetzten Verfahren waren zudem nicht auf Bewegungsdaten beschränkt und es hätten auch multimodale Klassifizierungen durchgeführt werden können, da diese von mobileSSI direkt unterstützt wurden.

Für die Akzeptanz eines derartigen Systems durch den Nutzer ist allerdings bedeutsam, dass dieser nicht mit zu vielen Fragen gestört wird. Hierfür könnten bei zukünftigen Systemen beispielsweise Verfahren wie „query by committee“ oder auch andere Sensoren, wie die Smartscale,

situationsbedingt eingesetzt werden. Eine kontextsensitive Auswahl der Ausgabemodalitäten und der Bestimmung eines passenden Zeitpunkts der Ausgabe könnten ebenfalls von Vorteil sein.

Mobiles interaktives maschinelles Lernen wurde auch im Bezug auf Erklärbarkeit (XAI) in der Arbeit (Flutura, [Seiderer](#), Huber, Weitz, Aslan, Schlagowski, André & Rathmann, 2020) eingesetzt. Hierbei wurden für die Klassifizierung von ästhetischem und unästhetischem Wald relevante Bildbereiche von Fotos hervorgehoben. Diese Markierungen konnten in einer Studie die Transparenz des maschinellen Lernens erhöhen.

8.3.6 Entwicklung eines adaptiven Roboterassistenten für Informationen über Getränke

Einleitung

Der Zuckergehalt vieler Lebensmittel ist unangemessen hoch, sodass die Gesundheit von Menschen geschädigt werden kann. Auch viele Fruchtsäfte besitzen von Natur aus hohe Zuckerkonzentrationen, die unverdünnt mit denen mancher Softdrinks vergleichbar sind. Vielen Menschen ist dies jedoch nicht direkt bewusst. Eine Lösung für dieses Problem kann hierbei ein Empfehlungssystem zur Bereitstellung von Informationen über Getränke darstellen. Insbesondere soziale Roboter oder Sprachassistenten können hierbei im häuslichen Bereich und teilweise auch in der Öffentlichkeit eingesetzt werden. Um dies mit einem sozialen Roboter zu untersuchen, wurde ein Prototyp entwickelt, der Angaben über die Menge an Kalorien sprachlich ausgeben konnte und zusätzlich versuchte, Nutzer zu überzeugen, eine gesündere Getränkewahl zu treffen. Hierzu sollte der Roboter selbstständig lernen, welche diesbezüglichen Sprachformulierungen am besten geeignet sind.

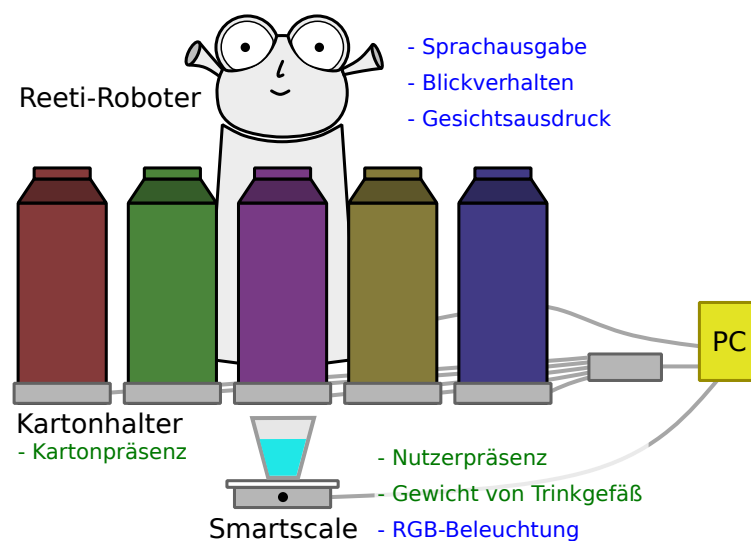


ABBILDUNG 8.21: Der „Drink-O-Mender“. Systemeingaben sind in grüner, Ausgaben in blauer Schriftfarbe.

Der entwickelte Prototyp zur Bereitstellung von Informationen zu Getränken ist in Abbildung 8.21 dargestellt. Er bestand aus einem Roboter (Reeti⁸), der auf einem Tisch platziert wurde sowie einer Smartscale, die die Getränkemenge (Wägezelle) und Präsenz eines Nutzers (Laser-Distanzsensoren) erfassen konnte. Zudem wurden speziell entwickelte Halter mit Sensorik für Getränkekartons mit verschiedenen Fruchtsäften und Eistee auf dem Tisch angeordnet. Ein PC / Laptop war für die Erfassung, Auswertung und Ausgabe zuständig. Diese Anordnung erhielt den Namen „Drink-O-Mender“. Das System wurde bei einer öffentlichen Wissenschaftsveranstaltung

⁸<http://www.reeti.fr/index.php/en/detailen> (13.01.2020)

(lange Nacht der Wissenschaft 2018 in Augsburg) evaluiert, weshalb es entsprechend robust gestaltet sein musste. Insgesamt interagierten innerhalb von vier Stunden 78 Nutzer mit dem System. In dieser Zeit konnte der Roboter seine Sprachformulierungen mittels Reinforcement Learning anpassen. Die Ergebnisse der Arbeit beinhalten Erkenntnisse über den Anpassungsvorgang sowie die Nutzerakzeptanz und Usability.

Die Inhalte dieses Abschnitts wurden in der eigenen Publikation (Ritschel, Seiderer, Janowski, Aslan & André, 2018) veröffentlicht.

Verwandte Arbeiten

Soziale Roboter und Ernährung Soziale Roboter wurden in der Literatur für die Abfrage der Kalorienanzahl von Lebensmitteln (siehe „Pillo“⁹ von Tao et al. (2016)), zum Erreichen von Diätzielen (Kidd & Breazeal, 2007) und Protokollierung der Kalorienaufnahme (Kidd & Breazeal, 2008) eingesetzt. Beim Einsatz von sozialen Robotern konnte bei der Protokollierung von Kalorien in einer Studie von Kidd und Breazeal (2008) eine erhöhte Langzeitmotivation festgestellt werden. Bei der Ausgabe von Ernährungsempfehlungen konnten zwischen Tablet-PCs und Robotern in Studien von Cruz-Maya und Tapus (2016) und Hammer et al. (2017) keine Unterschiede bezüglich der Überzeugungskraft ermittelt werden. Insbesondere von älteren Menschen wurde jedoch die Interaktion mit sozialen Robotern als weniger komplex und leichter erlernbar eingestuft (Hammer et al., 2017). Auch bei Kindern können soziale Roboter von Vorteil sein. So konnte in einer Studie von Short et al. (2014) mit einem „DragonBot“, der gesunde Ernährung an Kinder vermitteln sollte, ein erhöhtes Engagement beim Lernen und mehr Freude bei der Interaktion beobachtet werden.

Adaptive soziale Roboter Soziale Roboter, die ihr Verhalten an den Nutzer anpassen können, wurden in verschiedensten Szenarien genutzt. Bei diesem eigenen Projekt wurde Reinforcement Learning für die Anpassung des gesprochenen Inhalts eingesetzt. Reinforcement Learning wurde beispielsweise auch bei Robotern für das Erlernen des sozialen Verhaltens (Barraquand & Crowley, 2008), Lehren von Schülern (Gordon et al., 2016), Rehabilitation nach einem Schlaganfall (Tapus et al., 2008) sowie Intervention bei autistischen Kindern (Liu et al., 2008) angewandt. Reinforcement Learning wurde zudem verwendet, um die beste Strategie für ein dauerhaftes Nutzerengagement z. B. beim Schachspielen gegen den Roboter zu finden (Leite et al., 2011).

Wie bei den oben angeführten Beispielen wurde beim Drink-O-Mender ein adaptiver Ansatz genutzt. In diesem Fall sollte der Roboter Menschen hinsichtlich der Auswahl gesünderer Getränke unterstützen. Im Gegensatz zur Arbeit von Kidd und Breazeal (2007) lag der Fokus jedoch nicht ausschließlich auf der Nahrungserfassung, sondern auch darauf, den Inhalt der sprachlichen Ausgabe zu optimieren. Ein sozialer Roboter wurde als Interaktionspartner genutzt, da umfangreicheres soziales Verhalten vom Roboter ausgedrückt werden kann, die das menschliche Interesse verstärken (Kruijff-Korbayová et al., 2014) können. Zusätzlich könnten Roboter (verkörperte Agenten) Vorteile bei der Entscheidungsfindung besitzen (Shinozawa et al., 2005).

Implementierung

Soziale Roboter verfügen über verschiedene Formen der Kommunikation: Sprache, Blickverhalten, Gesichtsausdrücke oder Gesten. Um den Nutzer bei der Ernährung zu unterstützen, ist Sprache die wichtigste Ausgabeform. Beim entwickelten Prototyp wurde der gesprochene Inhalt über die Zeit, mittels vordefinierter Aussagen, abhängig von der Getränkewahl an eine Nutzergruppe, angepasst. In diesem Fall bestand die Nutzergruppe aus den Teilnehmern einer Wissenschaftsausstellung. Die Aussagen des Roboters betonten entweder die hohe Kalorienanzahl bei Getränken mit hohem Nährwert oder zogen die Aufmerksamkeit auf die mit den wenigsten Kalorien. Das Ziel war hierbei, gesundes Verhalten zu fördern, indem die Nutzer auf die Nährwerte aufmerksam gemacht

⁹<https://pillohealth.com/> (13.01.2020)

wurden. Diese Information war in einer Datenbank gespeichert. Die Kalorienzahl eines gewählten Getränks, ermittelt aus der Art des Getränks und der Menge, spielte auch bei der Berechnung der „Belohnung“ (engl. reward) für den Lernvorgang beim Reinforcement Learning eine Rolle und beeinflusste die Selektion der Aussagen. Der Mechanismus ist in der Veröffentlichung Ritschel et al. (2018) ausführlich beschrieben. Um die Erfassung der konsumierten Getränke zu erleichtern, wurde eine Smartscale entwickelt, mit der das Trinkgefäß mit Inhalt automatisch gewogen werden konnte, sodass keine explizite Angabe der Menge per Sprache oder Text notwendig war. Zusätzlich wurde das ausgewählte Getränk mit eigens entwickelten Getränkekartonhaltern erfasst. Beide Informationen dienten als „Belohnung“ im Adaptionsvorgang. Die Systemausgabe erfolgte mittels des oben bereits erwähnten Reeti-Roboters mit Sprachausgabe und nonverbalem Verhalten.

Hardware Abbildung 8.22 stellt im Überblick das System „Drink-O-Mender“ dar. Die wesentlichen Komponenten sind Smartscale, Gefäßhalter, PC bzw. Laptop und der Reeti-Roboter. Die Smartscale und Gefäßhalter waren hierbei per USB mit dem PC verbunden. Hierüber fand die Kommunikation per serieller Schnittstelle statt und die Komponenten wurden mit Strom versorgt. Der PC führte die komplette System-Logik aus. An den PC war der Reeti-Roboter per LAN angeschlossen, dessen Aktuatorik (Ton, Servomotoren, LEDs) über die Hersteller-API angesteuert wurde.

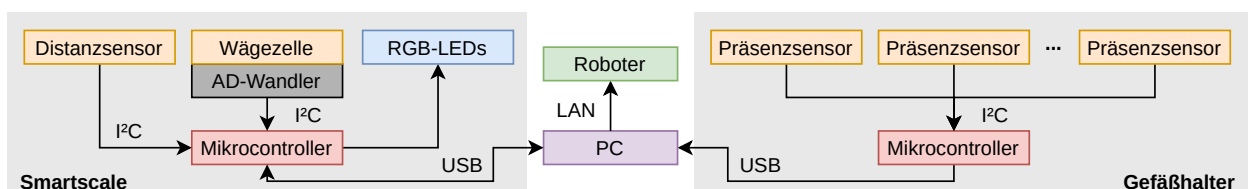


ABBILDUNG 8.22: Systemüberblick „Drink-O-Mender“.

Da die Getränkekartons vorgegeben (Position und Art) waren, konnte die Art des Getränks zuverlässig festgestellt werden, indem erkannt wurde, welcher der fünf Getränkekartons vom Nutzer genommen wurde. Die Positionen der Kartons mussten hierfür ebenfalls definiert werden. Hierfür gab es grundsätzlich verschiedene Lösungsansätze. Mittels Wägezellen oder Drucksensoren hätte erkannt werden können, welcher Getränkekarton angehoben wurde. Im einfachsten Fall wären hierfür auch Mikrotaster unter den Abstellflächen der Getränkekartons möglich gewesen. Für den Prototyp wurde allerdings eine optische Methode gewählt. Diese war kostengünstig, mechanisch stabil und konnte vor am Karton herunterlaufender Flüssigkeit leicht geschützt werden. Es wurde pro 3D-gedrucktem Getränkekartonhalter ein optischer Nahdistanzsensor vom Typ APDS-9960¹⁰ an der Rückseite verbaut. Abbildung 8.23 zeigt unten eine schematische Darstellung des Halters von der Seite. Diese Art von Sensor wird beispielsweise in vielen Smartphones verbaut, um während dem Telefonieren den Bildschirm abschalten zu können, wenn die obere Bildschirmseite des Geräts mit dem Sensor gegen ein Ohr gehalten wird. Der Sensor erfasst Objekte in direkter Nähe vor dem Sensor und ist gegenüber Sonnenlicht unempfindlich. Er kann per I²C-Schnittstelle mit Mikrocontrollern verbunden werden. Hierbei entstand allerdings ein Problem, da alle Breakout-Platinen mit dem SMD-Chip identische I²C-Adressen besaßen und diese nicht verändert werden konnten. Es war somit nicht möglich, an einen I²C-Bus mehrere Sensoren anzuschließen. Daher wurde ein I²C-Multiplexer vom Typ TCA9548A¹¹ von Adafruit eingesetzt, der zwischen angeschlossenen I²C-Geräten (mit identischer Adresse) umschalten kann. In diesem Fall wurden fünf verschiedene Sensoren benötigt, wobei dieser Multiplexer sogar für bis zu acht Geräte geeignet ist. Der Multiplexer wurde dabei vom Mikrocontroller zunächst per I²C so angesteuert, dass er auf

¹⁰<https://www.broadcom.com/products/optical-sensors/integrated-ambient-light-and-proximity-sensors/apds-9960> (10.02.2020)

¹¹<https://learn.adafruit.com/adafruit-tca9548a-1-to-8-i2c-multiplexer-breakout/overview> (10.02.2020)

einen bestimmten Sensor schaltete, bevor der Sensor vom Mikrocontroller angesprochen wurde. Als Mikrocontroller wurde ein Arduino Pro Micro (ATmega32U4) mit 3,3 V verwendet. Die einzelnen Sensoren in den Getränkekartonhaltern wurden mit fünfpoligen JST-Steckern mit der Platine verbunden, um einen einfacheren Transport zu ermöglichen. In Abbildung 8.24 rechts ist die Platine mit Mikrocontroller und I²C-Multiplexer sowie den Anschlüssen in einem 3D-gedruckten Gehäuse dargestellt. Die Firmware war so eingestellt, dass alle fünf Sensoren auf Objekte in einem Abstand von 3 cm reagierten und mit 10 Hz ausgelesen wurden. Sobald mindestens einer der Sensoren eine Zustandsänderung erfasste, wurde diese von der Firmware an den PC per serieller Schnittstelle übermittelt.

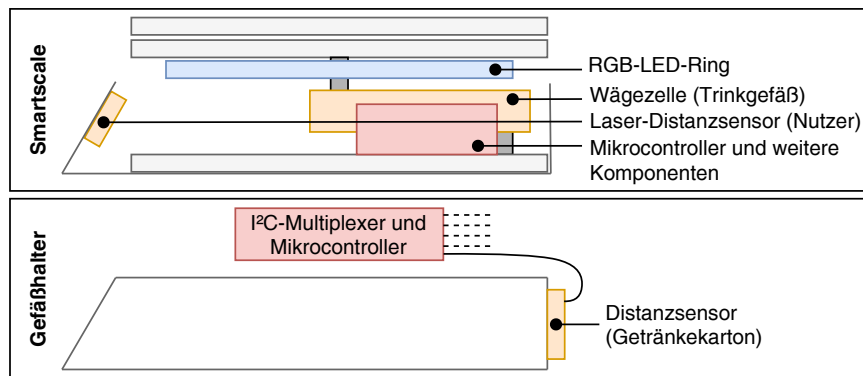


ABBILDUNG 8.23: Schematische Seitenansichten der Smartscale und eines Gefäßhalter-Prototyps.

Zur Bestimmung der Menge des getrunkenen Getränks wurde die bereits angesprochene Smartscale entwickelt (siehe Abbildung 8.23 oben und 8.24 links). Hierfür wurde eine Waagezelle mit einer Belastungsgrenze von 5 kg mit einem HX711-AD-Wandler mit I²C-Schnittstelle integriert. Eine automatische Kalibrierung der Waage fand beim Einschalten statt. Die Menge der Flüssigkeit konnte bestimmt werden, da das Gewicht der genutzten Trinkgefäße (Pappbecher mit einem Fassungsvermögen von 100 ml) bekannt war und bei der Dateninterpretation subtrahiert wurde. Zudem war im Prototyp ein ToF-Laser-Distanzsensor vom Typ VL53L0X integriert, der Abstände bis zu 200 cm vor dem Sensor erfassen konnte. Der Sensor verfügte ebenfalls über eine I²C-Schnittstelle. Der Abstandssensor ermöglichte dem System, Personen, die davor standen, zu erfassen. Hierfür wurde der Sensor in einem 30°-Winkel im 3D-gedruckten Gehäuse befestigt. Bei der Smartscale wurde unterhalb der zwei 3-mm-dicken, halbdurchsichtigen Acrylglasplatten ein Ring mit 24 einzeln ansteuerbaren RGB-LEDs vom Typ WS2812B eingebaut. Die beiden Acrylglasplatten streuten das Licht der LEDs, sodass eine diffuse Beleuchtung der Fläche möglich war, auf die das Trinkgefäß zum Wiegen gestellt werden sollte. Die obere Platte war mit transparentem Doppelklebeband mit der anderen verbunden und verdeckte zusätzlich die Schraube, die die untere Acrylglasplatte mit der Waagezelle verband. Aluminium-Winkel wurden an die Ränder der Acrylglasplatten geklebt und abgedichtet, sodass eventuell verschüttete kleinere Flüssigkeitsmengen, nicht direkt in die Elektronik fließen konnte. Verschiedene Farb-Helligkeits-Animationen waren bei den LEDs möglich, beispielsweise rotierende Farbverläufe. Um möglichst stufenlose Animationen zu ermöglichen, wurde ein leistungsfähiger ESP8266-Mikrocontroller eingesetzt, der neben der Ansteuerung der Sensoren, ausreichend schnell war, um die LEDs mit 30 Hz anzusteuern. Animationen konnten mittels serieller Schnittstelle ausgelöst und beendet werden. Der Mikrocontroller verfügte über ausreichend Speicher, um verschiedene Animationen zu speichern. Die Sensoren wurden ebenfalls mit 30 Hz ausgelesen und per serieller Schnittstelle an den PC übertragen. Die WLAN-Konnektivität des ESP8266 wurde in diesem Fall nicht genutzt, da das USB-Kabel auch für die Stromversorgung benötigt wurde und somit bereits eine stabile Kommunikationsmöglichkeit mit dem PC bestand. Die Stromstärke (500 mA) eines USB-2-Ports war für eine gut sichtbare Helligkeit der LEDs geeignet und wurde per Software limitiert, sodass die 500-mA-Grenze insgesamt

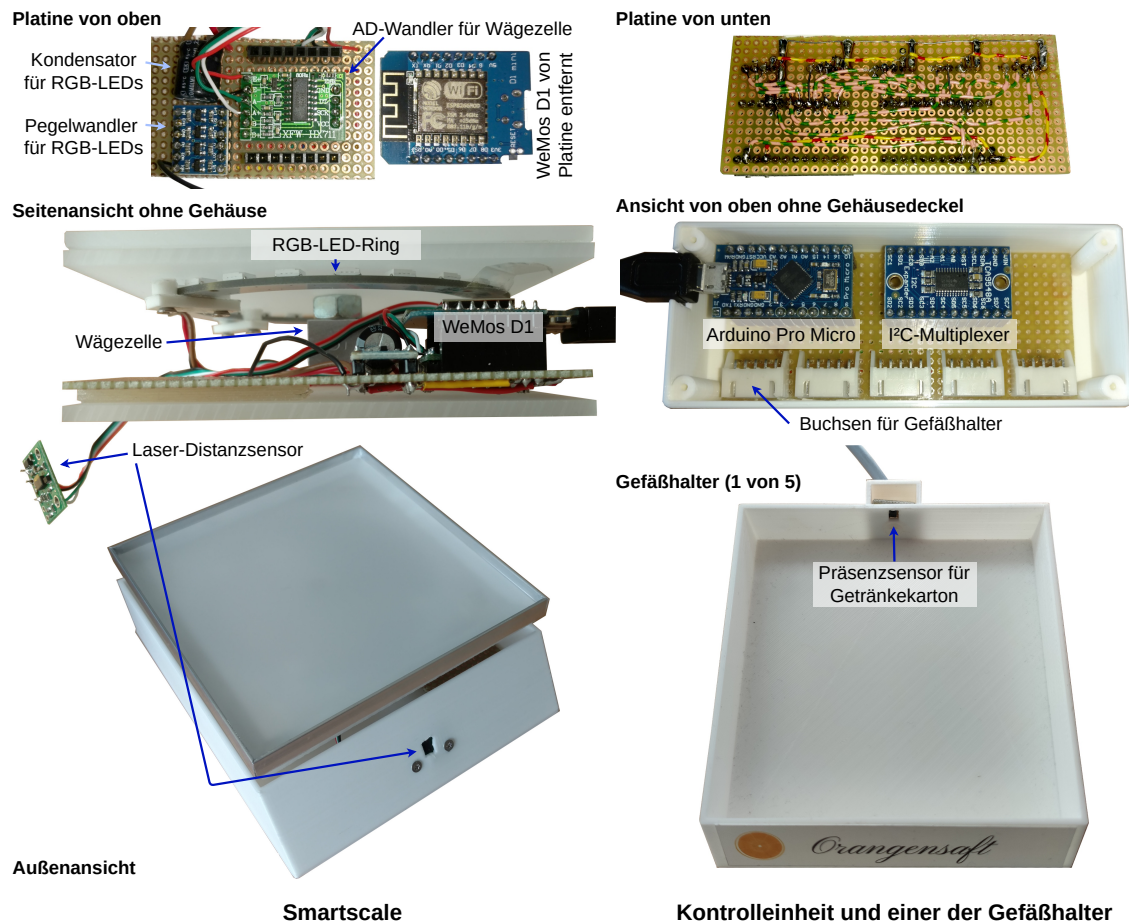


ABBILDUNG 8.24: Fotos von den eigenen Hardware-Prototypen. Links ist die Smartscale, rechts die Einheit mit dem Mikrocontroller und I2C-Multiplexer sowie einer von fünf Getränkekartonhaltern abgebildet.

nicht überschritten werden konnte. Um zu verhindern, dass ein Spannungsabfall beim Aktivieren der LEDs stattfindet, wurde ein Elektrolyt-Kondensator mit einer Kapazität von $1000\ \mu\text{F}$ parallel geschaltet. Der ESP8266-Mikrocontroller sowie die Sensorik, wurden mit $3,3\ \text{V}$ vom integrierten Spannungsregler der WeMos D1 kompatiblen Schaltung versorgt. Die WS2812B-LEDs benötigen allerdings eine Spannung von $5\ \text{V}$, die direkt vom USB-Port stammte. Für eine stabile Ansteuerung der einzelnen Controller der RGB-LEDs mit einem digitalen Ausgabepin des Mikrocontrollers wurde ein Signalpegelwandler eingesetzt, der den Logikpegel des Mikrocontrollers ($3,3\ \text{V}$) auf die benötigte TTL-Logik ($5\ \text{V}$) der LED-Controller umsetzte.

Signalverarbeitung Die Signalverarbeitung der Sensordaten wurde teilweise direkt von den Sensoren oder den Mikrocontrollern übernommen. Bei der Smartscale wurden die kalibrierten Gewichtswerte mit einem gleitenden Mittelwert über vier Werte bei einer Abtastrate von $30\ \text{Hz}$ von der Firmware des ESP8266-Controllers gefiltert. Bei den Gefäßhaltern wurde ein Schwellwert genutzt, um zu erkennen, ob ein Getränkekarton entfernt oder abgestellt wurde. Nur die Zustandsänderungen wurden in diesem Fall von der Firmware des ATmega32U4 an den PC weitergeleitet.

Der PC verwendete zur Datenverarbeitung die eigene Software „Eventerpretor“ (siehe Abschnitt 4.2.2). In diesem Fall wurden die Knoten zur Ansteuerung der seriellen Schnittstelle (Ein- und Ausgabe), gleitender Mittelwert, Schwellwert mit zeitlichen Verzögerungen und PostgreSQL-Datenbank (Ein- und Ausgabe) eingesetzt. Die Smartscale benötigte eine bidirektionale Kommunikation über die serielle Schnittstelle (Steuerung der LEDs / Auslesen der Sensordaten). Diese war

Nutzer seinen Trinkbecher auf diese gestellt hatte. Die Abstellfläche der Waage wurde mit einem rotierenden bunten Farbverlauf beleuchtet, um dem Nutzer zusätzlich zur Sprachanweisung und dem Blickverhalten zu signalisieren, dass das Trinkgefäß darauf abgestellt werden sollte. Während des Wiegevorgangs wurde die Farbe statisch auf Rot und zum Abschluss auf Grün geschaltet. Das ermittelte Gewicht wurde vom System gespeichert und zusätzlich verbal ausgegeben, wobei das (geringe) Eigengewicht des Pappbechers subtrahiert wurde. Anschließend wurde der Nutzer darauf hingewiesen, dass er nun trinken kann. Falls nach einer Pause noch nicht alle Getränkekartons zurückgestellt waren, wurde der Nutzer aufgefordert, diese aufzuräumen. Daraufhin wurde die Person verabschiedet und intern eine Aktualisierung des Lernvorgangs durchgeführt.

Beobachtungen bei der Systeminteraktion

Das System wurde bei der langen Nacht der Wissenschaft 2018 in Augsburg präsentiert (siehe Abbildung 8.26). Dieses System sollte einer breiteren Öffentlichkeit verfügbar gemacht werden, um Reaktionen hinsichtlich der Funktion des Drink-O-Menders zu erfassen und auszuwerten. Es handelte sich hierbei um eine öffentliche Wissenschaftsausstellung, bei der mehr als 2000 Besucher gezählt wurden. Das System war an einem Stand aufgebaut und mindestens zwei wissenschaftliche Mitarbeiter des Lehrstuhls waren anwesend, um nach der Systeminteraktion eines Besuchers Fragen beantworten zu können und Notizen anzufertigen. In diesem Kontext kann der Test des Systems nicht mit einer Laborstudie verglichen werden, da hierbei beispielsweise schlechte Lichtverhältnisse, hoher Geräuschpegel und teilweise Zeitmangel bei den Besuchern vorhanden waren. Das System war zudem hinsichtlich Umfang darauf ausgelegt, dass eine Interaktion nur relativ kurze Zeit erfordern würde. Es wurden daher nur Beobachtungen bei der Systeminteraktion notiert, die jedoch Anhaltspunkte über die Akzeptanz und die Usability des Systems über ein sehr heterogenes (Altersgruppen, Geschlecht, ...) Publikum geben.



ABBILDUNG 8.26: Fotos vom Systemaufbau bei der „langen Nacht der Wissenschaft“ in Augsburg 2018. (Bilder aus (Ritschel, Seiderer, Janowski, Aslan & André, 2018))

Für die Ausstellung wurden fünf Getränke mit verschiedenen Kaloriengehalten ausgewählt (siehe Abbildung 8.26): Tomatensaft und Eistee (geringster Kaloriengehalt), Apfel- und Orangensaft (mittlerer Kaloriengehalt) und Traubensaft (höchster Kaloriengehalt). In vier Stunden wurden 78 vollständige Interaktionen der Besucher mit dem System aufgezeichnet. Die Gesamtzahl der Interaktion war etwas höher, allerdings beendeten nicht alle Besucher die Interaktion oder es traten Inkonsistenzen auf, weil beispielsweise der Roboter durch laute Umgebungsgeräusche nicht verstanden wurde oder die Anweisungen des Roboters nicht eingehalten wurden. Teilweise versuchten ganze Personengruppen mit dem Roboter zu interagieren, das System konnte allerdings nur von einer Person genutzt werden.

Anpassung des Systems Der Anpassungsalgorithmus (Reinforcement Learning) verwendete den Kaloriengehalt des gewählten Getränks, um die „Belohnung“ für die ausgewählte Formulierung zu bestimmen. Desto höher dieser war, desto geringer war die Belohnung. Bei den 78 Systeminteraktionen wurden 21-mal Tomatensaft und Eistee ausgewählt, Apfel- und Orangensaft 25-mal und Traubensaft 32-mal. Offensichtlich war die vom Roboter gesprochene Information über den Kaloriengehalt nicht ausreichend, um Nutzer davon zu überzeugen, Getränke mit geringerer Kalorienanzahl zu wählen. Im Verlauf der Ausstellung favorisierte der Algorithmus den expliziten Hinweis auf den hohen Kaloriengehalt des Traubensaftes.

Das Experiment zeigte, dass die Hardware eine zuverlässige Quelle für die Bestimmung der Belohnung des Algorithmus darstellte. Der Algorithmus passte bei diesem System jedoch ausschließlich die Sprachformulierungen an. Eine automatische Anpassung der Argumentationen des Roboters im Hinblick auf eine gesunde Getränkeauswahl könnte in einem künftigen System die Überzeugungskraft des Roboters verstärken.

Irreführende Affordanzen Die meisten Besucher gingen davon aus, dass der Roboter über eine Spracherkennung verfügt. Eine solche wurde allerdings nicht implementiert, da eine hohe Lautstärke bei der Veranstaltung zu erwarten war und somit keine zuverlässige Interaktion möglich gewesen wäre. Besucher gingen höchstwahrscheinlich von einem Sprachinterface aus, da der Roboter Sprachausgabe nutzte. Die meisten Personen nahmen auch an, dass der Roboter die Kameras in den Augen nutzt, um ihr Gesicht zu erkennen bzw. zu erfassen, auf welches Getränk sie zeigen. Eine derartige Funktionalität wäre in zukünftigen Prototypen implementierbar. Auf Grund schlechter Lichtverhältnisse wäre dies bei der Wissenschaftsausstellung jedoch nur mit erheblichem Aufwand robust umsetzbar gewesen.

Einige Nutzer versuchten Getränkekartons auf der Smartscale abzustellen. Dies könnte darauf beruhen, dass durch die prominente Position und LED-Animation eine hohe Affordanz bestand, die Waage auf irgendeine Art einzusetzen. Der Trinkbecher wurde den Besuchern erst nach dem Nehmen eines Getränkekartons zum Befüllen übergeben, sodass dies zunächst das einzige Objekt war, das sie darauf stellen konnten. Die Anweisungen des Roboters bzw. der Ablauf sollten daher in zukünftigen Umsetzungen entsprechend angepasst werden.

Akzeptanz des Roboters Generell schienen die Besucher den Roboter zu mögen. Einige bezeichneten Reeti als „süß“. Mindestens ein Besucher zeigte Interesse daran, statt eine Smartphone-App zu nutzen, den Roboter auch zu Hause für Lebensmittelinformationen zu befragen. Andere zeigten sich weniger begeistert und würden es komisch finden, so etwas bei sich aufzubauen oder bezweifelten, dass jemand auf „ein Spielzeug“ hören würde. Ein anderer Teilnehmer war primär daran interessiert auszuprobieren, wie das System auf bestimmte Eingaben reagierte. Die letzten zwei Beobachtungen könnten durch Reetis eher harmloses Aussehen und unterwürfiges Verhalten ausgelöst worden sein. Um zu verhindern, dass Nutzer sich bevormundet fühlten, gab Reeti ausschließlich höfliche Vorschläge bezüglich einer gesunden Getränkeauswahl. Er kritisierte auch nicht, wenn der Nutzer diesbezüglich suboptimale Entscheidungen traf.

Um Nutzer zu motivieren, muss ein Roboter äußerlich ansprechend und im Auftreten sowie Empfehlungsverhalten zugleich überzeugend wirken. Neben der körperlichen Erscheinung ist es immer noch eine Herausforderung, die richtige Balance zwischen Höflichkeit und Überzeugungskraft zu finden, da diese von vielen situativen und individuellen Faktoren abhängen (Hammer et al., 2016). Dieser Aspekt könnte verbessert werden, wenn nicht nur der Inhalt der Empfehlungen, sondern auch deren Präsentation angepasst würden.

Einige Nutzer wünschten sich zudem mehr Informationen über die Getränke. Eine Verbesserung des Systems könnte daher auch darin bestehen, nicht nur die Kalorien, sondern z. B. zusätzlich den Vitamingehalt oder die Anbauweise der verarbeiteten Getränkezutaten zu berücksichtigen. Mittels eines Sprachinterfaces könnte der Roboter hierbei gezielt auf Fragen von Nutzern antworten.

Zusammenfassung

Das Experiment hat gezeigt, dass ein sozialer Roboter als einfaches Interface zur Informationsabfrage zu Nahrungsmitteln gut geeignet sein kann. Der präsentierte Prototyp hatte hierbei die Vorteile, dass er sein Verhalten während der Laufzeit anpassen konnte, leicht verwendbare, robuste Hardware nutzte und wenig Platz einnahm. Da auf Video- und Audioverarbeitung verzichtet wurde, könnte der Aufbau auch in Szenarien eingesetzt werden, in denen die Privatsphäre besonders geschützt werden muss.

Von den Beobachtungen lässt sich ableiten, dass der Roboter grundsätzlich akzeptiert wurde, jedoch teilweise zu wenig Autorität besaß. Beispielsweise wurde trotz der Information, dass der Traubensaft die meisten Kalorien beinhaltete, dieses Getränk von Nutzern am häufigsten gewählt. Für künftige Entwicklungen könnte dies beim Anpassungsvorgang berücksichtigt werden, indem auch eine noch überzeugendere Präsentation von Informationen einbezogen wird. Allerdings ist nicht auszuschließen, dass z. B. Traubensaft bei Kindern besonders beliebt ist und der Hinweis auf die hohen Kalorienwerte bewusst oder unbewusst nicht beachtet wurde. Gegebenenfalls sind auch persönliche Abneigungen bzw. Unverträglichkeiten gegenüber anderen Getränken maßgebend. Da vereinzelt jüngere Kinder mit dem System interagierten, könnten hierbei zudem Verständnisprobleme bezüglich der angegebenen Kalorienanzahl aufgetreten sein. Daher müssten besondere Zielgruppen wie Kinder vom System bezüglich der Informationsausgabe gesondert berücksichtigt werden.

Durch den eingesetzten Roboter und dessen Sprachausgaben erwarteten einige der Besucher, dass das System über eine Spracheingabe verfügt. Zukünftig könnte das System entsprechend erweitert werden. Eine Spracheingabe könnte jedoch nur bei geeigneter Geräuschkulisse eingesetzt werden und müsste die Privatsphäre der Nutzer respektieren.

8.3.7 Entwicklung eines deutschsprachigen Sprachassistenten für Senioren zur Abfrage von Informationen über Lebensmittel

Einleitung

Sprachassistenten werden immer beliebter, da sie in vielen Situationen die einfachste und intuitivste Methode darstellen, mit einem Computersystem zu interagieren. Insbesondere Senioren können von Sprachinterfaces profitieren, da sie einen leichteren Zugang zu modernen Assistenzsystemen ermöglichen. Senioren stellen jedoch eine besondere Zielgruppe dar, die altersbedingte Einschränkungen besitzt und häufig weniger vertraut mit aktueller Technik ist. Relativ viele Senioren verfügen derzeit zudem über keinen Internetanschluss bzw. nur eingeschränkten Internetzugang. Eine gute Internetanbindung ist allerdings für Sprachassistenten von kommerziellen Anbietern eine Mindestvoraussetzung, da diese den Großteil der Datenverarbeitung auf den Servern der Anbieter – in der Cloud – durchführen. Somit können schützenswerte Informationen von kommerziellen Onlineanbietern in unbekannter Form verwertet werden. Dies ist besonders kritisch, wenn Senioren gesundheitsbezogene Assistenzsysteme, wie beispielsweise für die Einnahme von Medikamenten oder Informationen zu Lebensmitteln nutzen möchten. Häufig ist diese Problematik Senioren nicht bekannt.

Ziel dieses Projekts war es demzufolge, einen Sprachassistenten zu entwickeln, der die Privatsphäre besonders beachtet und für deutschsprachige Senioren nutzbar ist. Der Prototyp sollte ausschließlich quelloffene Softwarekomponenten und maschinelles Lernen verwenden sowie vollständig offline lauffähig sein. Alle benötigten Daten sollten aus freien (open data) Datenbanken stammen. Als Anwendungsbeispiel wurde für den Sprachassistenten die Ausgabe von Informationen zu den Nährwerten von verpackten Lebensmitteln ausgewählt. Die Beschriftungen auf den Verpackungen sind häufig für Senioren nur mühsam mit Lesehilfen lesbar und verfügen zusätzlich nicht generell über gesundheitsbezogene Bewertungen der Menge bestimmter Nährstoffe wie z. B. Zucker. Beim Prototypen wurden die Daten von verpackten Lebensmitteln mittels Barcode

erfasst und mithilfe der „open food facts“-Datenbank¹² ausgewertet. Bei einem Workshop mit zwei Senioren wurde das System vorgestellt und semistrukturierte Interviews durchgeführt. Die Resultate werden im Anschluss diskutiert, sodass sie generell hilfreiche Informationen geben können, um Sprachassistenten für Senioren zu entwickeln.

Dieser Abschnitt basiert auf der eigenen Publikation (Seiderer, Ritschel & André, 2020).

Verwandte Arbeiten

Verständliche Nährwertangaben Bei diesem Projekt lag der Fokus darauf, Informationen darüber sprachlich auszugeben, wie gesund ausgewählte verpackte Lebensmittel sind. In Deutschland und einigen anderen Ländern existieren diesbezüglich derzeit noch keine einheitlichen, verpflichtenden Kennzeichnungen auf den Verpackungen, die eine schnelle Beurteilung ermöglichen. Dies sollte mit dem entwickelten Sprachassistenten vereinfacht werden, indem der „Nutri-Score“ ausgegeben werden kann. Der Nutri-Score (Julia et al., 2017) ermöglicht eine leicht verständliche Bewertung mit Buchstaben zwischen A und E. „A“ steht hierbei für ein gesundes und „E“ für ein ungesundes Lebensmittel. Diese Kennzeichnung wurde in einigen Ländern, wie z. B. Frankreich, bereits flächendeckend eingeführt und schien daher auch für den entwickelten, eigenen Prototyp gut geeignet zu sein.

Chat-Bots und Sprachassistenten In der Literatur können verschiedene Beispiele für Sprachassistenten und Chat-Bots gefunden werden, die einen Bezug zu gesunder Ernährung besitzen. In der Arbeit von Magnini et al. (2018) wurde das aufgabenorientierte Dialogsystem *CH1* für Zuckerkrankke vorgestellt. Es diente dazu, die aufgenommene Menge an Kohlenhydraten zu berechnen. Das System wurde für die italienische Sprache entwickelt, interpretierte Texteingaben und erzeugte textuelle Ausgaben. Ein Chat-Bot, der zu einer gesünderen Ernährung führen sollte, wurde von Casas et al. (2018) vorgestellt. Er verwendete den Chatfuel-Cloud-Dienst. Der Nutzer definierte hierbei seine eigenen Ernährungsziele, wie beispielsweise eine Reduktion des Fleischkonsums. Von Angara et al. (2017) stammte der Konversationsagent *Foodie Fooderson*, der dabei helfen sollte, Lebensmittelverschwendung zu reduzieren und die Essgewohnheiten der Nutzer zu verbessern. Er gab hierzu Rezeptempfehlungen, wobei Allergien oder Diäten berücksichtigt wurden. Das System basierte auf den *IBM Watson conversational services*.

Chat-Bots und Sprachassistenten für Senioren Weitere Arbeiten untersuchten Anwendung von Chat-Bots und Sprachassistenten mit der Zielgruppe Senioren. Das System *NESTORE* von Angelini et al. (2019) hatte das Ziel, Senioren dabei zu unterstützen, einen gesünderen Lebensstil zu erreichen. Es agierte als Trainer für verschiedene Gesundheitsgebiete, wie beispielsweise Ernährung. Eine der Möglichkeiten mit *NESTORE* zu interagieren bestand darin, einen Chat-Bot zu nutzen. Die Sprachverarbeitung nutzte hierbei das Rasa-Framework, das auch für den in dieser Untersuchung vorgestellten Prototyp, eingesetzt wurde. Jesús-Azabal et al. (2019) präsentierten eine Kombination aus Sprachassistent und Smartphone-App für ältere Menschen, um sie an ihre tägliche Medikamenteneinnahme zu erinnern. Ihr Assistent sollte offline funktionieren und für Senioren geeignete Sprache verwenden. Das System nutzte die Open-Source-Software *Snips* von Coucke et al. (2018) für den Sprachassistenten. Shalini et al. (2019) entwickelten einen Sprachassistenten für die Google-Assistant- und Amazon-Echo-Plattformen, um Gesundheitsdaten von Senioren zu verwalten. Zur Interaktion wurden kommerzielle Geräte mit Bildschirm eingesetzt. Das System wurde von Senioren sowie deren Familien genutzt und bot leicht verständliche, audiovisuelle Gesundheitsinformationen, die mit gesprochener Sprache angefragt werden konnten. In der Arbeit von Cheng et al. (2018) wurde ein Sprachassistent für ältere Menschen mit Typ-2-Diabetes entwickelt und evaluiert. Das System nutzte Google Home und ein Webinterface für Visualisierungen. Die Senioren in der Vorstudie präferierten dieses System gegenüber Smartphone-Apps.

¹²<https://de.openfoodfacts.org/> (31.01.2020)

Der Großteil der hier aufgeführten Arbeiten, die die Zielgruppe Senioren behandeln, wählten assistive und gesundheitsbezogene Aufgaben für ihre Chat-Bots bzw. Sprachassistenten. Die Interaktion mit den Systemen erfolgte meist mittels einer Kombination aus GUI und Sprachinterface. Diese Kombination wurde auch für den eigenen Prototyp eingesetzt.

Sprachassistenten und Komponenten

Sprachassistenten bestehen aus folgenden Komponenten: automatische Spracherkennung (engl. automatic speech recognition = ASR), natürliches Sprachverständnis (engl. natural language understanding = NLU), Dialogmanager (engl. dialog manager), natürliche Sprachgenerierung (engl. natural language generation = NLG) und Sprachausgabe (engl. text to speech = TTS). Derzeit verfügbare Lösungen von Amazon, Apple, Google oder Microsoft setzen primär auf cloudbasierte Systeme, die das Aufbauen eines Sprachassistenten oder Chat-Bots mit den Produkten des entsprechenden Anbieters vereinfachen. Die Nutzung von Online-Diensten für das Sammeln und Verarbeiten von Daten ist derzeit für viele Aufgaben verbreitet. Allerdings geht eine Datenverarbeitung mittels Cloud-Diensten immer mit dem Risiko der Verletzung der Privatsphäre einher. Es stellen sich hierbei auch die Fragen, woher die Informationen stammen oder mit welchen Firmen persönliche Daten geteilt wurden. Dies gilt insbesondere auch für die Spracherkennung, wenn die Audiodaten direkt an den Cloud-Anbieter gesendet werden.

Nachstehend werden zunächst in einem Überblick die Komponenten vorgestellt, die zur Entwicklung eines quelloffenen Sprachassistenten benötigt werden.

Frameworks für Sprachassistenten Es existieren einige Frameworks, bei denen alle notwendigen Komponenten für einen Sprachassistenten integriert wurden, und somit die Entwicklung vereinfachen. In Tabelle 8.4 ist eine Selektion quelloffener Frameworks angegeben (Stand: 2019/2020). Alle Frameworks unterstützten neben Englisch auch einige andere Sprachen, sind offline nutzbar und wurden aktiv weiterentwickelt. Hierbei sind die jeweils unterstützten Open-Source-Lösungen für die Spracherkennung und -ausgabe benannt.

Name	ASR	TTS
Jasper ¹³	Pocketsphinx, Julius	Festival TTS, Flite, pico TTS, MaryTTS
Kalliope ¹⁴	CMUSphinx	eSpeak, MaryTTS, pico TTS
Leon ¹⁵	DeepSpeech	Flite
LinTO ¹⁶	LinSTT (Kaldi based)	pico TTS
Mycroft ¹⁷	DeepSpeech or Kaldi	Mimic, eSpeak, MaryTTS
Open Assistant ¹⁸	CMUSphinx	eSpeak

TABELLE 8.4: Quelloffene Frameworks für Sprachassistenten.

Snips von Coucke et al. (2018) war ein populärer, quelloffener Sprachassistent, der das „privacy-by-design“-Prinzip verfolgte und über ein deutsches Sprachmodell verfügte. Diese Software hätte daher auch für die eigene Arbeit eine interessante Option dargestellt. Allerdings wurde nach einer Firmenübernahme eine wesentliche Komponente (Snips Console) des Systems deaktiviert. Sie lief auf dem Webserver der Firma und war augenscheinlich nicht quelloffen. Ohne diesen Systembestandteil wurde jedoch die Anpassbarkeit stark eingeschränkt. Daher wurde Snips in diesem Überblick nicht berücksichtigt.

¹³<https://jasperproject.github.io/> (13.01.2020)

¹⁴<https://kalliope-project.github.io/> (13.01.2020)

¹⁵<https://getleon.ai/> (13.01.2020)

¹⁶<https://linto.ai/> (13.01.2020)

¹⁷<https://mycroft.ai/> (13.01.2020)

¹⁸<https://openassistant.org/> (13.01.2020)

Derzeit leiden alle Lösungen für einen quelloffenen Sprachassistenten unter demselben Problem: die nicht-englische Sprachunterstützung der Spracherkennung und -ausgabe. Es ist häufig schwierig, ausreichend gut funktionierende deutsche Sprachmodelle zu finden. Ein moderner, auf tiefen neuronalen Netzen basierender Spracherkenner, ist DeepSpeech¹⁹ von Mozilla. Während der Entwicklung des Systemprototyps (Ende 2019) waren allerdings keine offiziellen, gut funktionierenden deutschen Sprachmodelle verfügbar. Die Sprachausgabe war beim entwickelten Sprachassistenten weniger kritisch, solange die Ausgabe verständlich war. Ein möglichst menschlich klingendes Modell wurde nicht benötigt. Daher waren hierfür auch ältere Lösungen nutzbar.

Frameworks für Chat-Bots Einige Lösungen für Sprachassistenten beinhalten bereits Komponenten für das Sprachverständnis (NLU) sowie Dialogmanagement. Diese sind allerdings häufig auch in Chat-Bot-Frameworks aufzufinden. Bei diesen fehlen lediglich die Spracherkennung (ASR) und Sprachausgabe (TTS). Eine solche Lösung ist Microsofts Bot Framework²⁰. Es ist quelloffen, allerdings wurde es darauf ausgerichtet, mit Microsofts Azure-Cloud-Diensten verwendet zu werden. Ein weiteres Framework ist Rasa von Bocklisch et al. (2017). Es beinhaltet natürliche Sprachverarbeitung (engl. natural language processing = NLP), die auch für nicht-englische Sprachen verfügbar ist sowie Dialogmanagement. Es existiert noch eine Vielzahl weiterer Frameworks, jedoch gehört Rasa zu den umfangreicheren Frameworks, das beispielsweise auch von Angelini et al. (2019) eingesetzt wurde.

Spracherkennung Tiefe neuronale Netze sind bei der Spracherkennung älteren Techniken, wie z. B. Hidden Markov Modellen (HMM), bei der Wortfehlerrate (engl. word error rate = WER) überlegen. Daher wurden nur Spracherkenner mit modernen Algorithmen berücksichtigt. Für zwei konnten deutsche Modelle gefunden werden. DeepSpeech fehlt derzeit ein offizielles deutsches Modell, jedoch wurden von Agarwal und Zesch (2019) Modelle erstellt und getestet. Diese Modelle waren während der Entwicklungsphase des Prototyps noch nicht verfügbar, daher wurde ein Modell für Kaldi (Povey et al., 2011) von Milde und Köhn (2018) eingesetzt.

Um den Nutzern volle Kontrolle zu erlauben, war eine automatische Aktivierung der Spracherkennung beim Prototyp nicht vorgesehen. Diese Funktionalität wäre allerdings möglich gewesen, wenn ein Open-Source-Schlüsselwort-Erkennen (engl. keyword spotter) zum Starten der Aufnahme und Sprachaktivitätserkennung (engl. voice activity detector = VAD) zum Stoppen der Aufnahme ergänzt worden wären. Zur Erkennung von Schlüsselwörtern hätte beispielsweise *mycroft-precise*²¹ eingesetzt werden können. Der Aufwand, ein ausreichend funktionales Modell für verschiedene Personen zu erstellen, sollte allerdings nicht unterschätzt werden, da hierbei auch Sprachassistenten von Großkonzernen häufig Probleme besitzen.

Sprachverständnis, Dialogmanagement und Sprachgenerierung Für die Sprachverarbeitung gab es drei populäre Lösungen mit deutschen Modellen: Apache OpenNLP²², Stanford CoreNLP von Manning et al. (2014) und SpaCy²³. SpaCy ist eine neuere Software und nutzt Deep-Learning-Methoden. Sie ist Teil des Sprachverständnisses von Rasa.

Für das Dialogmanagement existieren einige quelloffene Lösungen, die je nach Komplexität des Problems ausgewählt werden sollten. In diesem Fall wurde ein aufgabenorientiertes Dialogsystem benötigt. Eine Verarbeitung von komplexeren Dialogen, wie beispielsweise Smalltalk, war hierbei nicht vorgesehen. Daher war die in Rasa integrierte, vorlagenbasierte Lösung, ausreichend. Von Rasa wurde ebenso die Sprachgenerierung genutzt, wobei diese noch relativ begrenzt war.

¹⁹<https://github.com/mozilla/DeepSpeech> (03.02.2020)

²⁰<https://dev.botframework.com/> (05.02.2020)

²¹<https://github.com/MycroftAI/mycroft-precise> (05.02.2020)

²²<https://opennlp.apache.org/> (05.02.2020)

²³<https://spacy.io/> (05.02.2020)

Umfassendere Lösungen waren z. B. SimpleNLG von Gatt und Reiter (2009) oder RosaeNLG²⁴, die deutsche Sprachmodelle besitzen. Deep-Learning-Lösungen waren zur Entwicklungszeit in quelloffenen Lösungen nicht verbreitet.

Sprachsynthese Für die Sprachsynthese existierten einige ältere quelloffene Lösungen. Für den Prototyp wurde MaryTTS²⁵ eingesetzt, das hinreichend gute deutsche Sprachausgabe liefern konnte. Die erzeugten Audiodaten konnten zudem problemlos vom integrierten Server empfangen werden. Andere TTS-Software mit deutschen Modellen beinhaltet Tabelle 8.4. In Zukunft könnte Mozillas TTS²⁶ eine gute Lösung darstellen, da hierbei Deep-Learning-Methoden eingesetzt werden. Zusammen mit der Spracherkennung von Mozilla sollen diese in der Zukunft in Webbrowsern, als Teil der Web-Speech-API²⁷, eingesetzt werden.

Implementierung

Einen Überblick über das eigene, entwickelte System gibt Abbildung 8.27. Alle Systemkomponenten wurden in verschiedenen Docker-Containern in einer virtuellen Maschine (VM) installiert. Die virtuelle Maschine nutzte als Betriebssystem Xubuntu 19.04 x86-64. Der VM wurden acht physische CPU-Kerne (Intel Xeon E5-2690 v4), 8 GiB RAM und 80 GiB HDD-Speicher des Host-Systems zugeordnet. Die Leistung des Systems war daher vergleichbar mit einem Desktop-Rechner.

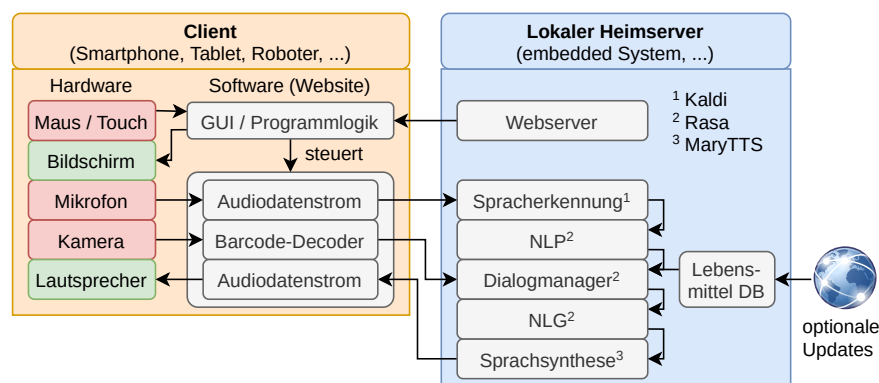


ABBILDUNG 8.27: Vereinfachter Systemüberblick.

Die Kommunikation zwischen den Softwarekomponenten wurde durch die Programmlogik auf dem Client ermöglicht. In der Abbildung ist dies vereinfacht dargestellt. Das GUI (siehe Abbildung 8.28) sowie die Programmlogik wurden mit einer HTML5-Webseite mit Javascript-Code implementiert.

Zunächst fotografierte ein Nutzer den Barcode eines Lebensmittels. Der Barcode wurde anschließend im Javascript-Code dekodiert, als Text ausgegeben und an das Rasa-Framework per WebSocket-Verbindung gesendet. Rasa sendete daraufhin einen generierten Text mit dem ermittelten Produktnamen zurück. Der Text wurde angezeigt und an MaryTTS gesendet. Die von MaryTTS erzeugte Sprachausgabe wurde vom Client empfangen und abgespielt. Nun konnte der Nutzer eine Frage stellen. Dies war entweder direkt per Text- oder mittels Spracheingabe möglich. Um die Spracherkennung zu nutzen, musste zunächst der Button „Aufnehmen“ betätigt werden. Daraufhin konnte der Nutzer den gewünschten Satz sprechen und zum Abschluss wiederum auf den Button drücken, um die Aufnahme zu beenden. Anschließend wurden die Audiodaten an den Spracherkennung gesendet. Der transkribierte Text wurde zurück an den Client geschickt und

²⁴<https://rosaenlg.org/> (05.02.2020)

²⁵<http://mary.dfki.de/> (05.02.2020)

²⁶<https://github.com/mozilla/TTS> (05.02.2020)

²⁷<https://wicg.github.io/speech-api/> (05.02.2020)

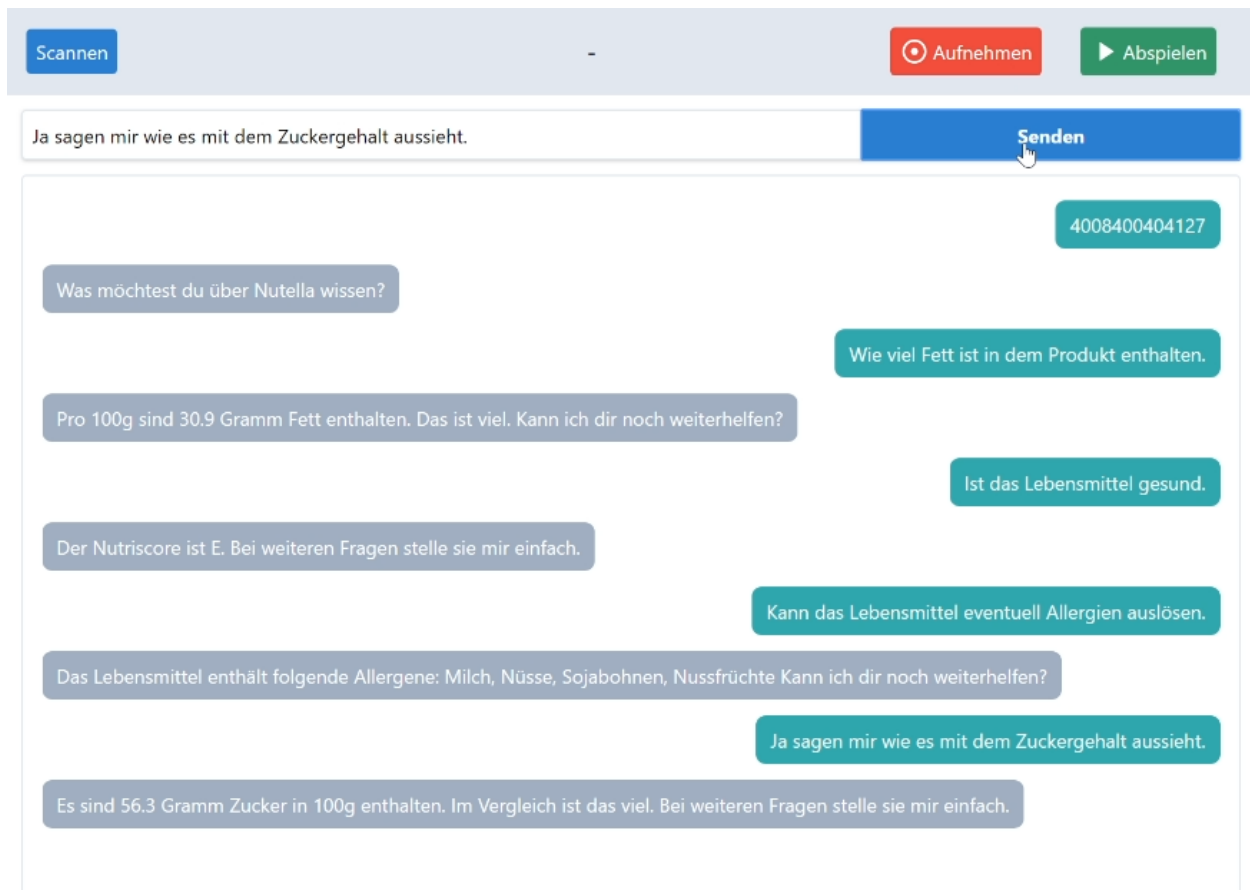


ABBILDUNG 8.28: Exemplarischer Dialog. Der Barcode am Anfang wurde vom Nutzer mit der Bilderkennung eingescannt (mittels des Buttons „Scannen“). Nutzereingaben sind türkis hinterlegt. Die Texte der Spracherkennung wurden hier nicht angepasst oder korrigiert. Die Spracheingabe erfolgt mit dem „Aufnehmen“-Button. Eine manuelle Texteingabe oder Korrektur des erkannten Textes ist möglich. Die Sprachausgabe kann mit „Abspielen“ wiederholt werden.

vom GUI dargestellt. Der Text wurde zugleich an Rasa zur Sprachverarbeitung gesendet, das eine Textausgabe generierte und an den Client sandte. Dieser Text wurde vom GUI angezeigt und an die TTS-Engine gesendet, deren Audioausgabe vom Client abgespielt wurde.

Für die Spracherkennung wurde Kaldi mit dem deutschen Sprachmodell *tuda_swc_mailabs_voc400k*²⁸ eingesetzt. Dies war das leistungsfähigste, allerdings auch größte Modell, das von Milde und Köhn (2018) bereitgestellt wurde. Für die Integration wurde der vorgeschlagene Kaldi GStreamer-Server eingesetzt, an den unkomprimierte Mono-Audiodaten mit einer Abtastrate von 16 kHz per WebSocket-Verbindung gesendet werden konnten. Für einen Satz wie „Sage mir wie viel Zucker in dem Lebensmittel enthalten ist.“ benötigte die Spracherkennung zwischen dem Absenden der Audiodaten und Empfang des erkannten Texts ca. vier Sekunden. Dies stellte eine akzeptable Verzögerung dar, da die anderen Systemkomponenten praktisch kaum wahrnehmbare Latenzen besaßen. Die Geschwindigkeit des Spracherkenners hätte jedoch erhöht werden können, indem beispielsweise ein kleineres Modell (mit schlechterer Erkennungsrate) verwendet oder die GPU-Beschleunigung aktiviert worden wäre. Es sollte auch beachtet werden, dass das Modell hinsichtlich des Vokabulars nicht auf diese Aufgabe optimiert wurde. Die Erkennungsgenauigkeit war bei den durchgeführten Tests mit den integrierten Mikrofonen eines Laptops bzw. Smartphones in relativ leiser Umgebung akzeptabel. In diesem Fall war es allerdings auch nicht nötig, dass alle Wörter korrekt transkribiert werden, da die Sprachverarbeitung nicht alle Wörter berücksichtigen

²⁸<https://github.com/uhh-lt/kaldi-tuda-de> (05.02.2020)

musste. Optional konnte der Text auch per Tastatur eingegeben werden. Dies wäre hilfreich, falls beispielsweise in lauten Umgebungen die Spracherkennung nicht funktioniert hätte oder eine Spracheingabe unerwünscht gewesen wäre.

Das Sprachverständnis, Dialogmanagement und die Sprachgenerierung wurden vom Rasa-Framework bereitgestellt. Das System konnte Informationen über enthaltene Nährstoffe, wie den Zuckergehalt, Kalorien, Kohlenhydrate und Proteine in Gramm pro 100 g des Lebensmittels zurückgeben. Für eine erleichterte Interpretation des Zucker- und Fettgehalts wurden diese kurz erklärt, indem angegeben wurde, ob es sich um eine geringe, mittlere oder große Menge handelt. Hierzu wurden, wie auch bei *open food facts*, die Grenzwerte von Hagen (2010) genutzt. Zusätzlich konnten der Nutri-Score und enthaltene Allergene ausgegeben werden. Die bei Rasa als „Intents“ bezeichneten Nutzerabsichten wurden von der Sprachverarbeitung erkannt, um Anfragen des Nutzers zu erkennen. Eine spezifische Information, wie z. B. Zucker werden bei Rasa als „Entity“ (Entität) bezeichnet. Auf Basis von insgesamt 43 Beispielsätzen konnte das Sprachverständnis die verschiedenen „Intents“ und „Entities“ erlernen. Als Basismodell für SpaCy wurde *de_core_news_sm-2.1.0* verwendet. Um auf Nutzereingaben reagieren zu können, die das System nicht versteht, wurde ein zusätzlicher „Intent“ angelegt. In solchen Fällen konnte das System nochmals nachfragen. Die für die Sprachgenerierung benötigten Daten wurden mittels sogenannter „Actions“ (Aktionen) ermittelt. Diese bestanden aus Python-Funktionen, die in diesem Fall die Informationen aus einer lokalen SQLite-Datenbank abfragten. Sie enthielt die aktuellsten Daten von *open food facts*. Es wurden allerdings Einträge entfernt, bei denen benötigte Informationen fehlten. Die Datenbank umfasste 1.110.884 Produkte. Aktuellere Versionen der Datenbank konnten jederzeit über das Internet eingespielt werden. Durch die Nutzung der lokalen Datenbank konnten geringere Latenzen erreicht werden. Zudem war es für die Anbieter der Datenbank nicht möglich, festzustellen, über welche Produkte ein Nutzer Informationen abgefragt hat.

Als Sprachausgabe wurde MaryTTS mit dem deutschen Modell *dfki-pavoque-styles* genutzt. Es erzeugte meistens akzeptable Ergebnisse. Insbesondere Produktnamen und deren Herstellernamen konnten jedoch Probleme bei der Generierung verursachen, wobei die Sprache in solchen Fällen schwer verständlich war. Diese Informationen dienten allerdings nur der Kontrolle, ob das System den Barcode korrekt erkannt hatte, und sie konnten auch im GUI nachgelesen werden.

Evaluation

Die Implementierung zeigte, dass es bereits technisch möglich ist, einen deutschen Sprachassistenten basierend auf quelloffener Software zu implementieren. Dieser Prototyp wurde während eines eigenen Workshops mit Senioren eingesetzt.

Der Workshop wurde von einem Studienleiter durchgeführt, der zunächst den Studienablauf und die Verwendung der aufgezeichneten Daten (schriftlich) erklärte. Anschließend wurden die Senioren nach ihrem Alter und ihre Erfahrungen mit Sprachassistenten, Smarthomes und Smartphones gefragt. Danach wurde der Prototyp den Senioren vorgeführt und sie konnten diesen ausprobieren. Abschließend wurde ein semistrukturiertes Interview mit beiden Teilnehmern durchgeführt. Der Workshop fand in der Wohnung der Senioren statt und dauerte insgesamt ca. 60 Minuten.

Für die Studie wurde ein älteres Ehepaar (eine Frau und ein Mann) mit einem Alter zwischen ca. 60 und 70 Jahren (Mittelwert: 67,5) ausgewählt. Sie lebten unabhängig in ihrer eigenen Wohnung. Sie hatten bereits etwas Erfahrung mit einem installierten Smarthome-System, das allerdings primär Informationen ausgab und Hausautomation durchführte. Der Zugriff auf das System erfolgte hierbei per Webbrowser. Sie verfügten zudem über ein Smartphone, das sie allerdings nur gelegentlich nutzten. Beide Studienteilnehmer besaßen keine Erfahrungen mit Sprachassistenten für Smarthomes oder Smartphones. Die Technologie war ihnen allerdings aus Fernsehsendungen, Werbung und Nachrichten bekannt.

Der Sprachassistent wurde den Senioren vom Studienleiter vorgestellt. Für die Demonstration wurden jeweils ein Trinkkarton mit Orangen- bzw. Grapefruitsaft ausgewählt, die in der lokalen Lebensmitteldatenbank enthalten waren. Ein Microsoft Surface 4 Pro (Intel m3 CPU) Tablet mit Windows 10 wurde für die Darstellung der Website (Frontend) eingesetzt. Die Backend-VM lief auf einem kompakten PC mit AMD Ryzen-5-3600-CPU. Beide Geräte befanden sich im selben Netzwerk. Anhand des Orangensaft-Kartons wurde den Senioren erklärt, wie mit dem System interagiert werden konnte und welche Informationen abgefragt werden konnten. Anschließend testeten die Senioren das System eigenständig mit der Packung Grapefruitsaft.

Zum Abschluss wurde ein semistrukturiertes Interview durchgeführt. Es wurde gefragt, ob Probleme auftraten oder ob ihnen Verbesserungen bzw. Erweiterungen für den Prototyp einfallen. Zusätzlich wurde nach den Situationen und Orten gefragt, an denen die Teilnehmer sich vorstellen konnten, einen solchen Sprachassistenten für Nahrungsinformationen zu nutzen. Die Akzeptanz eines Smartphones oder Smartdisplays als Interaktionsgeräte wurde ebenfalls erfragt. Die Wahrung der Privatsphäre spielte bei der Entwicklung des Prototyps eine große Rolle, daher wurden die Senioren gefragt, ob sie zustimmen würden, Daten (insbesondere Audio) mit kommerziellen Firmen zu teilen, wenn sie hierfür von einer besseren Spracherkennung, Sprachausgabe und größere Funktionalität profitieren könnten. Zudem war von Interesse, welche Nährwerte für die Senioren relevant waren und ob der Nutri-Score und die Bewertung des Zuckergehalts für sie hilfreich waren.

Ergebnisse und Diskussion

Technischer Aufbau Während der Evaluation funktionierte die Spracherkennung in einer ruhigen Umgebung hinreichend gut. Sie erkannte die meisten der Fragen der Senioren. ASR hängt allerdings stark von der Audiohardware und Signalverarbeitung ab. Es existieren auch optimierte Open-Hardware- / Open-Source-Plattformen, wie beispielsweise ReSpeaker²⁹ für Raspberry Pi. Diese könnten bei zukünftigen Prototypen die Leistung der ASR verbessern.

Während des Testens des Prototyps trat ein Problem auf, das bereits während der Entwicklung auftrat und daher bekannt war. Das spezielle Wort „Nutri-Score“ wurde vom Spracherkennung nicht erkannt. Stattdessen versuchte die ASR, das Wort mit ähnlichen Begriffen zu ersetzen. Um dieses Problem zu beheben, wäre es nötig gewesen, ein neues Sprachmodell mit geeigneten Sprachsamples zu trainieren. Für ein generisches Modell wäre dies allerdings nur mit großem Aufwand möglich gewesen. Das Problem wurde umgangen, indem der Nutri-Score ausgegeben wurde, wenn danach gefragt wurde, ob das Lebensmittel gesund war oder welche Bewertung es besaß. Da das Sprachverständnis nicht alle Wörter benötigte, konnte das System auch trotz teilweise falsch erkannter Wörter die Frage eines Nutzers verstehen. Um die Spracherkennung robuster zu gestalten, wäre eine domänenspezifische ASR von Vorteil gewesen, wie sie auch bei *Snips* (Coucke et al., 2018) verwendet wurde. Das Problem mit unbekannten Wörtern trat auch bei der Sprachausgabe aus. Insbesondere der Marken- oder Firmenname eines Lebensmittels konnte falsch ausgesprochen werden. In den meisten Fällen konnten die Studienteilnehmer diese allerdings dennoch verstehen.

Die Kombination aus grafischem Touch- und Sprachinterface war für den Prototyp geeignet. Es war möglich zu sehen, was das System verstanden hatte und ein Verlauf bereits gestellter Fragen sowie der Antworten war erkennbar. Die Möglichkeit die Audioaufnahme mit dem GUI zu starten und stoppen erlaubte eine größere Kontrolle über das System. Zudem war das GUI sehr hilfreich bei der korrekten Positionierung des Barcodes eines Lebensmittels vor der Kamera, da es das aktuelle Kamerabild anzeigte. In zukünftigen Prototypen wäre es zudem möglich, Visualisierungen bezüglich der Lebensmittel über das Display anzuzeigen, wie es beispielsweise in der Arbeit von Shalini et al. (2019) der Fall war.

²⁹<https://github.com/respeaker/seeed-voicecard> (28.04.2020)

Interaktion von Senioren mit Sprachassistenten Senioren besitzen verschiedene altersbedingte und teilweise auch gesundheitsbedingte Einschränkungen, die bei der Entwicklung der Interaktion mit einem Sprachassistenten berücksichtigt werden müssen. Ihr Seh- und Hörvermögen sowie die Motorik sind häufig reduziert. Siehe hierzu auch die in Abschnitt 9.1.1 detaillierter ausgeführten allgemeinen, körperlichen Einschränkungen dieser Zielgruppe. Auf Lebensmittelverpackungen sind zwar bereits Nährwertangaben angegeben, die Studienteilnehmer bestätigten allerdings, dass sie die abgedruckten Texte öfters nur schwer lesen können, wenn die Schrift zu klein oder der Kontrast zu gering sind. Daher benötigten sie in solchen Fällen eine Lesehilfe und gute Beleuchtung. Da diese Informationen vom Systemprototyp ausgegeben wurden, wäre das System auch in solchen Situationen hilfreich für sie. Die Senioren konnten sich vorstellen, einen Sprachassistenten in der Küche zu verwenden, um dort Informationen darüber zu erhalten, wie gesund ein Lebensmittel ist.

Um eine gute Usability zu gewährleisten, wurden die GUI-Elemente und Texte vergleichsweise groß dargestellt. Generell kann die Interaktion mittels Touchscreen bei Senioren vorteilhaft sein, da es gemäß Motti et al. (2013) die Hürden senken kann, mit einem Computersystem zu interagieren. Um die Verwendung des Systems zu vereinfachen, waren mit der GUI jedoch nur wenige Interaktionen notwendig. Um akustische Verständnisprobleme zu vermeiden, wurde die Lautstärke der Sprachausgabe so eingestellt, dass sie von den Studienteilnehmern gut verstanden werden konnten. Der kapazitive Touchscreen des eingesetzten Tablets war hinreichend empfindlich, um die Eingaben der Senioren zu erfassen.

Ein aktueller Überblick über Sprachinterfaces für Senioren ist mit der Arbeit von Stigall et al. (2019) verfügbar. Wie die Autoren allerdings feststellten, gibt es vergleichsweise wenige Untersuchungen bei denen das Alter der Nutzer bei der Entwicklung von sprachbasierten Interfaces berücksichtigt wurde. In der Arbeit von Kobayashi et al. (2019) wurde untersucht, welche Auswirkungen altersbedingte kognitive Einschränkungen auf die Nutzung von Sprachinterfaces besitzen. Beispielsweise wird von den Autoren erwähnt, dass Senioren teilweise längere Pausen zwischen den Wörtern verwenden. L. Werner et al. (2019) stellten zudem fest, dass abhängig vom Alter Unterschiede bei der *Word Error Rate* von Spracherkennern auftreten. Es könnte daher nötig sein, diese Unterschiede bei der Modellerstellung zu berücksichtigen, damit eine Spracherkennung für Senioren vergleichbar gut funktioniert.

Privatsphäre Die Nutzung eines Privatsphäre schützenden Ansatzes für den Sprachassistenten wurde von den Senioren begrüßt. Insbesondere hätten sie das Teilen ihrer Sprachdaten mit Cloud-Diensten missbilligt, da sie dies bereits bei ihrem Smart TV, Smartphone und PC mit Windows 10 zu verhindern versuchten. Probleme bezüglich der Privatsphäre traten bei kommerziellen Produkten bereits des Öfteren auf (Hoy, 2018) und diese waren den Studienteilnehmern teilweise bekannt. Tatsächlich stellte auch in einer Studie von McLean und Osei-Frimpong (2019) mit 724 Sprachassistenten-Nutzern das wahrgenommene Risiko für die Privatsphäre einen wesentlichen Grund dar, Sprachassistenten nicht zu verwenden. Die beiden Senioren würden eine schlechtere Leistung eines Sprachassistenten eher akzeptieren, als die Daten mit einem kommerziellen Anbieter zu teilen. Lediglich die Nutzung einer kommerziellen Lösung, die vollständig offline arbeitet, könnte einen Kompromiss darstellen.

Lebensmittelinformationen und Erweiterungen Die Studienteilnehmer merkten an, dass sie die Ausgaben des Systems bezüglich der Nährwerte komfortabler empfanden, als es auf der Packung nachlesen zu müssen. Auch die zusätzlichen Interpretationen der Nährwerte des Systems mittels Nutri-Score und Bewertung des Zuckergehaltes wurden von den Senioren positiv aufgenommen. Zusätzlich hätten sie sich allerdings noch gewünscht, Informationen über die enthaltenen Mineralstoffe und Vitamine zu erhalten. Diese Art von Information wird allerdings von Produzenten normalerweise nicht angegeben und könnte nur teilweise aus speziellen Datenbanken, wie dem

„Bundeslebensmittelschlüssel“³⁰ erhalten werden. Diese Datenbanken sind allerdings in Deutschland nicht frei zugänglich. Eine nützliche Erweiterung könnte laut den Senioren eine Kombination mit einer Lebensmittelverwaltung darstellen, die sie beispielsweise darüber informieren könnte, wenn ein Lebensmittel in näherer Zukunft abläuft oder ein Produkt zur Neige geht.

Zusammenfassung

In diesem Abschnitt wurde ein Überblick über quelloffene Frameworks für Sprachassistenten sowie deren Komponenten gegeben. Zusätzliche Anforderungen waren hierbei, dass sie offline lauffähig waren und neben Englisch, insbesondere auch Deutsch unterstützten. Des Weiteren wurde die Implementierung eines eigenentwickelten, aufgabenorientierten Sprachassistenten als Anwendungsfall beschrieben. Dieser konnte Fragen über Nährstoffe eines verpackten Lebensmittels beantworten, indem der Barcode sowie eine lokale Datenbank verwendet wurden. Der Prototyp zeigt, dass aktuelle quelloffene Lösungen genutzt werden können, um ein sprachbasiertes, privatsphäreschützendes Interface zur Informationsabfrage bereitzustellen, ohne sich auf Cloud-Dienste abstützen zu müssen.

Im Rahmen eines Workshops wurde eine Evaluation des entwickelten Prototyps mit zwei Senioren durchgeführt. Hierbei zeigte sich, dass insbesondere ältere Personen von einem Sprachassistenten profitieren können, wenn deren besonderen Bedürfnisse berücksichtigt und ihre Privatsphäre geschützt werden.

8.4 Zusammenfassung des Kapitels

In diesem Kapitel wurden Möglichkeiten zur Erfassung von Nahrungsmitteln untersucht. Zunächst waren dies zwei automatische **Klassifikationsverfahren**. Bei der Klassifikation mit Audiodaten konnte anhand des Geräusches beim Essen und Sprechen verschiedene Nahrungsarten unterschieden werden. Bei den Gaskonzentrationen war es mit hohen Erkennungsraten möglich, Lebensmittelarten zu unterscheiden, wenn sich diese in einem geschlossenen Behälter befanden. Es wurden für die Datenaufzeichnung selbstentwickelte Prototypen eingesetzt. Da es sich um kostengünstige, unkalibrierte, einzigartige Gassensorik handelte, wurde zusätzlich ein Verfahren mit **Transferlernen** untersucht und erfolgreich angewandt, um die Klassifikationsmodelle für andere Sensoren desselben Typs nutzen zu können. Anschließend wurde ein System zur Erleichterung der manuellen Erfassung von Lebensmittelarten und Mengen entwickelt und evaluiert, das eine Smartscale (eigene Prototypen), Smartwatch und ein Smartphone kombinierte, um deren jeweiligen vorteilhaften Eigenschaften auszuschöpfen. Es konnte in einer Studie gezeigt werden, dass eine solche **Gerätekombination** potenziell den Aufwand bei der Erfassung von Lebensmitteln reduzieren und die Präzision erhöhen kann. In einer weiteren Untersuchung wurde **interaktives maschinelles Lernen** eingesetzt, um die Vorteile der Kombination aus automatischer Erfassung und manuellen Eingaben vom Nutzer auszuschöpfen. In den Experimenten konnte ein Klassifikator zur Erkennung einer Trinkbewegung, basierend auf Beschleunigungsdaten einer Smartwatch, durch interaktives maschinelles Lernen verbessert werden. Kooperatives maschinelles Lernen wurde in diesem Projekt zur Erstellung des Warmstartmodells eingesetzt, das als Basis für das interaktive maschinelle Lernen diente.

Anschließend wurden zwei Untersuchungen bezüglich der Ausgabe von Informationen zu Lebensmitteln durchgeführt. Zunächst wurde ein System, basierend auf einem Reeti-**Roboter**, einer Smartscale (eigener Prototyp) und Getränkekartonsensoren (eigene Prototypen) untersucht. Dieses konnte, basierend auf der Flüssigkeitsmenge und der Art des Getränks, Informationen darüber geben, wie viel Zucker das Getränk enthielt und somit, wie gesund es war. Der Prototyp musste hinreichend robust gestaltet werden, sodass er während einer Ausstellung mit vielen Besuchern

³⁰<https://www.blisdb.de/blis?background> (28.04.2020)

präsentiert und eingesetzt werden konnte. Abschließend wurde die Entwicklung eines quelloffenen, deutschen **Sprachassistenten** zur Abfrage und Bewertung von Nährwerten verpackter Lebensmittel vorgestellt. Dieses System wurde mit Senioren in einem Workshop als Beispiel herangezogen, um Probleme zu finden, die ältere Menschen mit Sprachassistenten haben können.

Die in dieser Arbeit untersuchten Methoden zur Erfassung von Lebensmitteln sind prinzipiell geeignet, um die Protokollierung von Lebensmitteln zu erleichtern. Hierfür wurden diverse mobile Geräte sowie Sensorik und Hardware aus dem IoT-Bereich eingesetzt. In der Forschung ist das Problem der automatischen Nahrungsmittelerfassung weiterhin ungelöst. Vor allem, da hierfür keine Sensorik existiert, die in allen Situationen die Menge und Art von aufgenommenen Lebensmitteln zuverlässig bestimmen könnte. Systeme, die auf unterschiedlichen Arten von Sensoren und Geräten basieren, können daher eine Lösung darstellen. Eine Erfassung kann zusätzlich mit maschinellem Lernen erleichtert werden. Insbesondere der Einsatz von Transferlernen und interaktivem maschinellem Lernen können hierbei zu verbesserten Erkennungsraten führen.

Soziale Roboter und Sprachassistenten, die kontextbezogene Nahrungsinformationen oder Empfehlungen geben, stellen für Anwender attraktive Lösungen dar. Vor allem Sprachassistenten erfahren eine steigende Verbreitung im Smarthome, aber auch bei mobilen Geräten. Insbesondere Senioren, die häufig eine geringere Technikaffinität besitzen, könnten von einer automatischen Erfassung und natürlichsprachigen Sprachinterfaces profitieren. Hierbei sollte allerdings auch der Schutz der Privatsphäre nicht vernachlässigt werden.

9 Projekt CARE: Verbesserung des Wohlbefindens von Senioren

In den vorherigen Kapiteln wurden bereits verschiedene Verfahren und Systeme vorgestellt, untersucht sowie Lösungsansätze aufgezeigt, die einen relevanten Beitrag zur Erfassung und Verbesserung des menschlichen Wohlbefindens leisten können. Der Schwerpunkt dieses Kapitels liegt auf der Beschreibung eines umfassenderen Systems, das alle wesentlichen Kategorien des Wohlbefindensmodells (siehe Abschnitt 2) abdeckte und speziell für die Bedürfnisse von Senioren entwickelt wurde. Das AAL-System (Ambient Assisted Living) sollte, basierend auf einem sensorgestützten Computersystem, kontextsensitive Empfehlungen bereitstellen, um typische Alltagsprobleme zu lösen und das Wohlbefinden sowie die Lebensqualität von Senioren positiv zu beeinflussen. Das System wurde im Rahmen des BMBF-Projektes „CARE“¹ (a sentient Context-Aware Recommender system for the Elderly) (01.01.2014 - 30.01.2016) an der Universität und Hochschule Augsburg entwickelt und evaluiert. Die im Projekt definierten Forschungsaufgaben wurden zunächst in Kooperation mit der National Technical University of Athens untersucht und nach Beendigung des BMBF-Projektes an der Universität und Hochschule Augsburg weiterverfolgt.

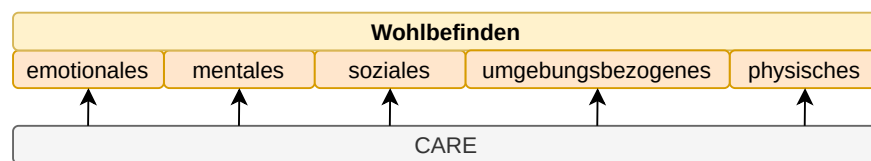


ABBILDUNG 9.1: Überblick über den Zusammenhang des CARE-Projekts sowie weiterführender Arbeiten in diesem Kapitel und dem Wohlbefindensmodell.

Das kontextsensitive Empfehlungssystem berücksichtigte alle Kategorien des Wohlbefindensmodells. Dies gilt ebenso für die weiterführenden Arbeiten in diesem Kapitel (vereinfachter Überblick siehe Abbildung 9.1). Die benötigten Kontextinformationen wurden automatisch bestimmt oder abgefragt. Das System wurde mit der Zielgruppe nutzerzentriert entwickelt und durchlief hierbei drei Iterationen des Entwicklungsprozesses (siehe Abbildung 9.2), die in diesem Kapitel beschrieben werden. In Abschnitt 9.4 wird CARE-Prototyp 1 und in Abschnitt 9.5 werden die CARE-Prototypen 1.5 und 2 beschrieben.

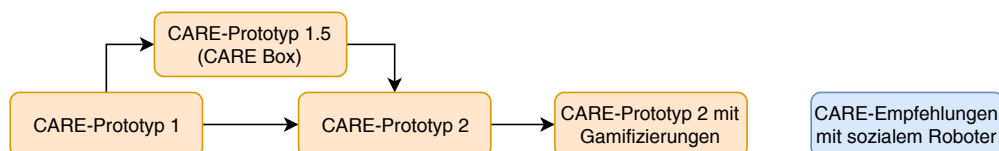


ABBILDUNG 9.2: Ablauf der Entwicklung der verschiedenen CARE-Prototypen (orange) und davon abgeleiteten Untersuchungen (blau).

Um das Interesse der Senioren am Empfehlungssystem nachhaltiger gewährleisten zu können, wurden Spielelemente (Gamifizierungen) in eine Weiterentwicklung des Prototyps 2 integriert und deren Wirkung untersucht (siehe Abschnitt 9.6). Dies war auch Ziel für die Entwicklung eines

¹<http://care-project.net>

AAL-Systems mit einem sozialen Roboter und speziellem Bedienpanel. Erkenntnisse vorheriger Prototypen flossen in diese Entwicklung ein (siehe Abschnitt 9.7).

Inhalte dieses Kapitels basieren auf folgenden Arbeiten mit eigener Beteiligung: (Hammer, Seiderer, André, Rist, Kastrinaki, Hondrou, Raouzaïou, Karpouzis & Kollias, 2015; Herpich, Rist, Seiderer & André, 2017; Rist, Seiderer & André, 2018; Rist, Seiderer, Hammer, Mayr & André, 2015; Ritschel, Janowski, Seiderer, Wagner & André, 2019; Ritschel, Seiderer, Janowski, Wagner & André, 2019; Seiderer, Hammer, André, Mayr & Rist, 2015).

9.1 Konzept CARE-System

Die Grundidee für das CARE-System bestand darin, ein hoch personalisiertes Assistenzsystem zur Unterstützung von (alleinlebenden) Senioren im häuslichen Wohnumfeld zu entwickeln. Hierbei sollten kontextsensitive Empfehlungen, primär auf der Basis von Sensordatenanalyse / -interpretation, generiert und zeitgerecht über Displays präsentiert werden. Dies sollte so erfolgen, dass sich die Menschen in ihrer Wohnumgebung nicht gestört fühlen. Das System sollte einen Bilderrahmen- und Empfehlungs-Modus besitzen, wobei im Bilderrahmen-Modus Bilder aus einer Foto-Sammlung ambient angezeigt werden sollten. Wenn sich ein Nutzer direkt vor dem System befand, sollten zum Kontext passende Empfehlungen zur Verbesserung des Wohlbefindens präsentiert werden. In Abbildung 9.3 ist das Konzept des CARE-Systems, gemäß dem Projektantrag, dargestellt.

Zunächst wurde in diesem Rahmen eine Anforderungsanalyse durchgeführt, indem die Bedürfnisse der Zielgruppe genauer betrachtet wurden, Befragungen stattfanden und Literatur ausgewertet wurde.

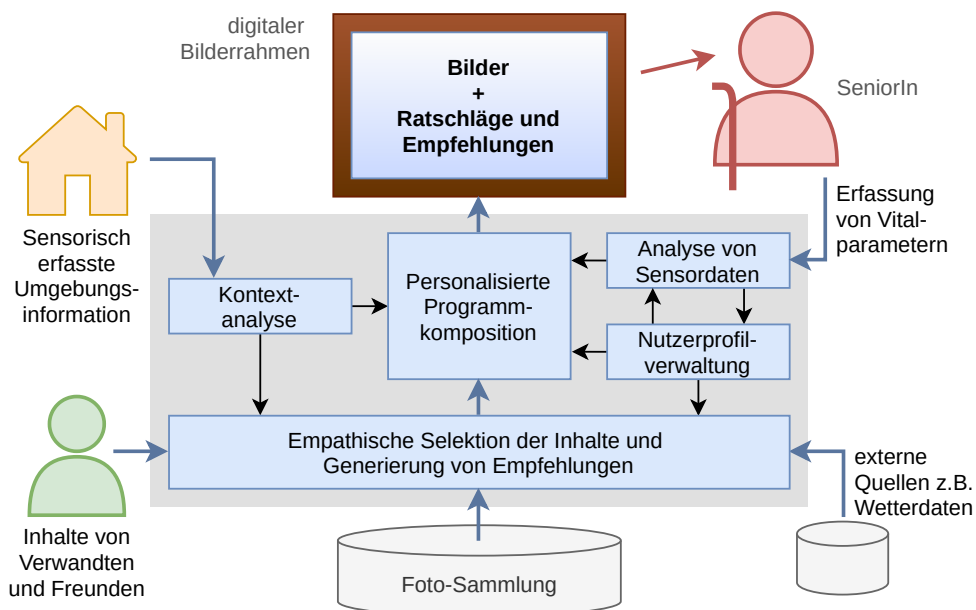


ABBILDUNG 9.3: Konzeptueller Aufbau des CARE-Systems (gemäß Projektantrag).

9.1.1 Anforderungen der Zielgruppe Senioren

Das CARE-System wurde speziell für die Zielgruppe Senioren entwickelt. Als Senioren werden hierbei Menschen, die älter als ca. 65 Jahre sind, bezeichnet. Die Zielgruppe musste im nutzer-zentrierten Entwicklungsprozess berücksichtigt werden. Es war zu beachten, dass, wie bei jeder Altersgruppe, individuelle Unterschiede zwischen Personen existieren, die beispielsweise krankheitsbedingt sind. Zudem können auch altersbedingte körperliche und kognitive Einschränkungen

sowie psychologische Probleme ihren Alltag so stark beeinflussen, dass sie nicht mehr ohne fremde Hilfe leben können. Für die in dieser Arbeit entwickelten Systeme wurde allerdings davon ausgegangen, dass die älteren Personen eigenständig leben oder nur in geringem Umfang externe Hilfe von Pflegediensten oder Familienmitgliedern benötigen.

Einschränkungen der Wahrnehmung, Kognition, Motorik und Hautleitfähigkeit

Generelle Einschränkungen bei der Wahrnehmung im Alter betreffen alle Sinne: sehen, hören, riechen, schmecken und tasten. Insbesondere das Hör- und Sehvermögen spielte bei Interaktionen mit den entwickelten Systemen eine maßgebliche Rolle. Des Weiteren waren kognitive und motorische Einschränkungen zu berücksichtigen. Im folgenden Abschnitt werden, wenn nicht anderweitig angegeben, Inhalte der Autoren Fisk et al. (2009) zusammengefasst.

Junge Erwachsene können Audiofrequenzen bis zu 15 kHz wahrnehmen. Ab einem Alter von ca. 65 / 70 Jahren treten **auditive** Limitierungen auf, sodass Frequenzen oberhalb von 4 kHz unhörbar werden können. Frequenzen unterhalb von 1 kHz sind hierbei nicht betroffen. Diese altersbedingte Veränderung beim Hören wird Presbyakusis genannt. Lautstärken von 8 dB bis 130 dB sind normalerweise wahrnehmbar, wobei eine normale Konversation bei ca. 60 dB liegt. Falls eine Abweichung der Hörschwelle von ca. 35 dB überschritten wird, liegt eine Schwerhörigkeit vor. Schwerhörigkeit ist nicht zwangsläufig altersbedingt, tritt hier allerdings häufiger auf und sollte daher berücksichtigt werden. Gemäß Jayakody et al. (2018) sollte zudem beachtet werden, dass Schwerhörigkeit neben beispielsweise sozialer Isolation auch zu kognitiven Einschränkungen führen kann. Bei der Entwicklung von Systemen, die Sprachausgabe nutzen, sollte auf eine höhere Lautstärke geachtet und lautere Umgebungsgeräusche möglichst vermieden werden.

Visuelle Einschränkungen sind ebenfalls nicht ausschließlich auf das Alter zurückzuführen, allerdings besteht auch hier mit zunehmendem Alter eine höhere Wahrscheinlichkeit. Eine Korrektur des Sehvermögens ist bei den meisten Senioren über 65 zufriedenstellend möglich, allerdings ist die Nutzung von Gleitsichtbrillen nicht immer optimal. In der Veröffentlichung werden weitere Einschränkungen beschrieben, wie Dunkelheitsadaption, Blendungsempfindlichkeit, Sichtweite und Verarbeitungsgeschwindigkeit der Wahrnehmung. Beim Design von visuellen Nutzerschnittstellen können beispielsweise eine größere Schrift und höhere Kontraste die Lesbarkeit für Senioren verbessern.

Die **kognitiven** Fähigkeiten von Senioren sind nur teilweise beeinträchtigt. Beim Arbeitsgedächtnis sowie dem semantischen, prozeduralen und prospektiven Langzeitgedächtnis treten unterschiedliche Arten der Einschränkung auf. Im Arbeitsgedächtnis werden relativ wenige, aktuelle Informationen verarbeitet. Bei einer Reduktion der kognitiven Leistungsfähigkeit kann beim Sprachverständnis der Inhalt gesprochener Sätze nur langsam verarbeitet werden, sodass wiederholtes Nachfragen ein Indikator für ein reduziertes Arbeitsgedächtnis sein kann. Beim Langzeitgedächtnis sind geringere Verschlechterungen feststellbar. Beim semantischen Gedächtnis, das beispielsweise faktische Informationen speichert, tritt eine Reduktion der Zugriffsgeschwindigkeit auf; es kann zudem das Abrufen von Wissen erschwert sein. Im prospektiven Gedächtnis sind Informationen über Aktivitäten gespeichert, die in der Zukunft durchgeführt werden sollen. Falls Erinnerungen oder Hinweise vorhanden sind, können im Alter allerdings keine größeren Einschränkungen festgestellt werden. Beim prozeduralen Gedächtnis tritt eine Einschränkung in der Art auf, dass neue Prozeduren langsamer und weniger erfolgreich gelernt werden können. Bei der Entwicklung von Systemen ist es daher von Vorteil, möglichst auf bereits bekannte Erfahrungen zurückzugreifen und das Systemverhalten vorhersehbar zu gestalten. Dies erlaubt auch älteren Menschen eine schnellere und weniger anstrengende Interaktion.

Die Aufmerksamkeit unterliegt nur teilweise altersbedingten Veränderungen. Die selektive (z. B. die Suche nach einer visuellen Ausgabe) sowie dynamische (Neuorientierung des Aufmerksamkeitsfokus) Aufmerksamkeit sind reduziert. Altersbedingte Unterschiede bezüglich der Informationsverarbeitung vergrößern sich, wenn die Komplexität von Aufgaben zunimmt. Zudem

ist eine Verringerung der Leistungsfähigkeit bei der Koordination mehrerer Aufgaben zu beobachten. Bei Interfaces kann eine gezielte Lenkung der Aufmerksamkeit beispielsweise auf einen bestimmten Bereich einer Anzeige eingesetzt werden.

Beim Sprachverständnis von Senioren können keine Einschränkungen festgestellt werden, wenn sie ihr semantisches Gedächtnis nutzen können. Falls allerdings Schlussfolgerungen notwendig werden und das Arbeitsgedächtnis überlastet ist, kann das Sprachverständnis reduziert werden. Bei der Nutzung von Sprachausgaben können daher zusätzliche Leitlinien berücksichtigt werden. Für ältere Menschen ist es beispielsweise leichter Wörter zu erkennen, wenn ihnen der Kontext bekannt ist. Zusätzlich können ein klarer Satzbau und deutliche Pausen an wichtigen Stellen die Verständlichkeit verbessern. Die Sprechgeschwindigkeit sollte hierbei nicht zu hoch sein. Für das Verständnis textueller Angaben ist ebenfalls zu beachten, dass die Sätze einen klaren Aufbau besitzen und, dass Senioren etwas mehr Zeit benötigen können, diese zu verstehen.

Bei den **Bewegungen** von Senioren ist zu beobachten, dass diese generell 1,5- bis 2-mal langsamer reagieren als jüngere Menschen und die Aktionen häufig unpräziser und unbeständiger sind. Dies ist auch bei der Computerinteraktion von Bedeutung. Die größeren Latenzen und Ungenauigkeiten müssen beispielsweise bei Eingabegesten (z. B. doppeltes Antippen) und der Größe von Buttons bei GUIs beachtet werden.

Senioren besitzen gemäß dem Fachbuch über Hautalterung von Quan, 2016, S. 173 häufiger sehr trockene Haut (Xerosis), wodurch deren **Leitfähigkeit** reduziert ist. Dies kann dazu führen, dass Geräte mit kapazitiver Sensorik, wie beispielsweise moderne Touchscreens mobiler Geräte, schlecht auf Fingereingaben reagieren. Die Empfindlichkeit der Sensorik ist gerätespezifisch und hängt unter anderem von der Herstellerkalibrierung ab. Dieses Problem betrifft allerdings alle Menschen mit trockener Haut oder dicker Hornhaut an den Fingerspitzen und ist auch unter dem englischen Begriff „Zombie Finger“² bekannt. Allerdings kann die Nutzung von Feuchtigkeitscreme dieses Problem reduzieren.

Befragung der Zielgruppe

In der eigenen Forschungsarbeit (Hammer, **Seiderer**, André, Rist, Kastrinaki, Hondrou, Raouzaïou, Karpouzis & Kollias, 2015) wurden die Anforderungen für den ersten CARE-Prototyp gesammelt. Hierzu wurden mit zwei Gruppen von Senioren strukturierte Interviews durchgeführt, die Fragen zum Lebensstil, den medizinischen Bedürfnissen, der Einstellung bezüglich AAL-Technologien sowie zur erwünschten Funktionalität und Systemkonfiguration eines Empfehlungssystems enthielten. Eine Seniorengruppe bestand aus 20 griechischen Personen (Alter zwischen 65 und 88; 12 Frauen, 8 Männer), die andere aus 27 deutschen Senioren (Alter zwischen 59 und 92; 15 Frauen, 12 Männer). Ungefähr die Hälfte der deutschen Seniorengruppe lebte alleine, bei der griechischen Gruppe etwas mehr als die Hälfte (13 Personen). Die Interviews bezogen sich bei den deutschen Teilnehmern darauf, wie Personen mit dem Empfehlungssystem generell interagieren könnten. Bei der griechischen Studie standen die Modalitäten zur Gewinnung der Aufmerksamkeit für die Empfehlungen im Vordergrund.

Im Folgenden werden die relevanten Ergebnisse der Interviews zusammengefasst. Die Senioren aus der deutschen und griechischen Gruppe zeigten unterschiedliche Einstellungen gegenüber AAL-Technologie. Gemeinsam war, dass sie sich damit sicherer fühlen würden oder die Familie entlasten könnten. Gleichzeitig waren sie allerdings auch besorgt, dass ihre Autonomie und Privatsphäre beeinträchtigt werden könnte. Nahezu kein Teilnehmer der griechischen Gruppe hätte eine Kamera in der Wohnung akzeptiert. Generell zeigten die Ergebnisse eine kritischere Einstellung gegenüber AAL-Technologien, als beispielsweise in der früheren Studie von Steinke et al. (2012) festgestellt wurde, bei der deutsche Senioren großes Vertrauen in AAL-Technologien zeigten. Es ist hierbei allerdings zu beachten, dass die Senioren durchaus Vorteile in der Verwendung von AAL-Systemen sahen, sie allerdings erst in späteren Lebensphasen für sie interessant wären oder

²<https://www.consumerreports.org/cro/news/2015/06/zombie-finger-and-touchscreens/index.htm> (22.06.2020)

für andere Senioren, die weniger fit sind. Die Senioren könnten sich stigmatisiert gefühlt haben, Technologie zu nutzen, die speziell für ältere Personen entwickelt wurde. Dieses Problem sollte bei der Entwicklung eines Gesundheits-Empfehlungssystems für Senioren berücksichtigt werden. Empfehlungen des Systems sollten eine Bevormundung vermeiden und die Privatsphäre respektieren. Die am häufigsten gewünschten Leistungen eines Empfehlungssystems waren:

- Erinnerungen an die Medikamenteneinnahme, regelmäßiges Trinken und wichtige Termine,
- Notrufe an Familienmitglieder und Mediziner im Fall von Unfällen und Krankheit,
- Warnungen bezüglich nicht abgeschalteter elektrischer Geräte, Lichter oder offener Fenster beim Verlassen des Hauses oder vor dem zu Bett gehen.

Zusätzlich sahen Senioren Vorteile darin, Empfehlungen für gesunde Ernährung, körperliche Aktivitäten und mentales Training zu erhalten. Deutsche Senioren empfanden Empfehlungen hilfreich, die sich positiv auf ihr soziales Leben auswirken, während die griechischen Senioren daran weniger Interesse zeigten.

Hinsichtlich der Ausgabemodalität konnte bei den griechischen Senioren festgestellt werden, dass diese visuelle Darstellungen gegenüber auditiver Ausgabe bevorzugten. Die Ausgabe der Empfehlungen sollte nach dem Wunsch der deutschen Senioren personalisiert zu passenden Zeitpunkten erfolgen. Dies galt auch für die griechischen Senioren, die sogar befürchteten, dass sie Empfehlungen verpassen könnten.

Im Hinblick auf die Systemkonfiguration bestand bei beiden Studien grundsätzliches Interesse an mobilen Geräten. Die deutsche Studie zeigte keine eindeutige Bevorzugung einzelner Interaktionsgeräte, wie Tablet, digitaler Bilderrahmen, Smartphone und Smartwatch. Smartwatches wurden jedoch teilweise geringfügig bevorzugt. Allerdings bezweifelten die Senioren in beiden Ländern, vor allem wegen der kleinen Bildschirme, dass sie die Geräte gut bedienen können. Diese Probleme konnten bereits in den Untersuchungen von Leung et al. (2012) bestätigt werden. Es zeigte sich hierbei zusätzlich, dass es vielen Senioren schwerfällt, die Anwendung und praktische Nutzung von Mobiltelefonen zu erlernen, die Geräte regelmäßig aufzuladen und mitzuführen.

Bei einem Einsatz robotischer Begleiter, die in der deutschen Studie betrachtet wurde, erwarteten die Senioren, dass die Roboter über umfassende Aktivitäten und Leistungen verfügen. Dies wurde ebenfalls in der Arbeit von Vardoulakis et al. (2012) festgestellt. Insbesondere die Hilfe im Haushalt, z. B. beim Tragen von schweren Objekten, schien bei den Senioren in der eigenen Studie eine Grundbedingung für die Akzeptanz von Robotern zu sein.

9.1.2 Verwandte Literatur

Ausgabe bei Empfehlungssystemen zur Verbesserung des Wohlbefindens

Es existierten bereits einige Systeme, die das Wohlbefinden ihrer Nutzer verbessern sollten, indem sie Daten in einer leicht verständlichen Art aufbereiteten. Eine Möglichkeit ist hierbei die Verwendung von **Metaphern**. Bei der mobilen Anwendung „BeWell“ von Lane et al. (2014) wurden beispielsweise körperliche Aktivitäten und Schlafmuster auf verschiedene Lebewesen eines virtuellen Aquariums übersetzt. Das mobile UbiFit-System von Consolvo et al. (2008) verfolgte einen ähnlichen Ansatz, indem körperliche Aktivität mittels der Metapher eines virtuellen Gartens vermittelt wurden.

Des Weiteren können die Daten in **natürliche Sprache** übersetzt werden. Diese Ausgabeform wurde beispielsweise von der Smartphone-App „Health Mashups“ von Bentley et al. (2013) umgesetzt und half dabei, die Beziehungen zwischen Lebensstil und Wohlbefinden leichter zu erkennen.

Das Motivate-System von Y. Lin et al. (2011) setzte Gesundheitsdaten des Nutzers und den Kontext in situationsbezogene **Empfehlungen** um. Hierbei wurde davon ausgegangen, dass Nutzer den Empfehlungen eher folgen, falls die Vorschläge zur richtigen Zeit und am passenden Ort gegeben werden.

Studien mit virtuellen **Agenten**, die Ratschläge zum Lebensstil und Gesundheit gaben, wurden beispielsweise von Yasavur et al. (2014) sowie Ring et al. (2015) durchgeführt. Ein verkörperter Agent (Roboter) wurde von Turunen et al. (2008) eingesetzt. Es gibt Hinweise darauf, dass Senioren positiv auf solche Begleiter reagieren, insbesondere falls sie soziales Verhalten nachahmen. Vardoulakis et al. (2012) merkten allerdings an, dass Senioren eine Vielzahl an sozialen Aktivitäten wie Smalltalk oder Kartenspielen erwarten, was eine zeit- und technisch aufwändige Implementierung voraussetzt.

Digitale Bilderrahmen

Digitale Bilderrahmen wurden in der Literatur bereits für einige Systeme, die das Wohlbefinden in verschiedener Weise verbessern sollten, verwendet.

Für Büroarbeiter wurden digitale Bilderrahmen von Obermair et al. (2008) eingesetzt, um deren Sitzverhalten zu verbessern. Die Bilderrahmen zeigten hierbei animierte Porträts von bekannten oder geliebten Menschen, die Feedback bezüglich des Verhaltens gaben.

Digitale Bilderrahmen wurden auch für ältere Menschen in verschiedenen Arbeiten bereits eingesetzt, um innerhalb von Familien die Beziehung zu den Senioren zu stärken. Consolvo et al. (2004) sowie Rowan und Mynatt (2005) nutzten hierfür ein Porträt der älteren Person, das entsprechend Daten über die Gesundheit und das Wohlbefinden des Seniors verändert wurde. Lindley et al. (2009) nutzten in ihren Untersuchungen Bilderrahmen zur Kommunikation mit Familienmitgliedern. Cornejo et al. (2013) erweiterten diesen Ansatz, indem sie mittels eines interaktiven Displays Senioren ermöglichten, soziale Exergames mit Familienmitgliedern zu spielen. Als Exergame wird eine Kombination aus Übung (engl. exercise) und Spiel (engl. game) bezeichnet.

Interaktion von Senioren mit mobilen Geräten

Die in diesem Kapitel bereits angesprochenen meist altersbedingten Einschränkungen und spezielle Präferenzen von Senioren wurden bei der Entwicklung der Prototypen berücksichtigt. In der Literatur sind hierzu einige praktische Erfahrungen bei der Entwicklung von Systemen für Senioren zu finden.

Obwohl Mobilgeräte für Senioren sehr hilfreich sein können, bestehen bei diesen eine Vielzahl an Usability-Problemen, da sie nicht an die speziellen Anforderungen von Senioren angepasst wurden. In der Arbeit von Angelini et al. (2013) wurde beispielsweise ein smartes Armband explizit für Senioren entwickelt, da herkömmliche Geräte für diese nicht geeignet waren. Die Autoren gingen allerdings auch davon aus, dass künftige Senioren, die bereits mehr Kontakt mit Computertechnologie besaßen, wiederum andere Anforderungen an die Geräte stellen werden.

Im Literaturüberblick und der Feldstudie von Leung et al. (2012) zeigte sich, dass Senioren (ab 65 Jahren) generell Probleme haben, die Nutzung mobiler Geräte zu erlernen. Gemäß den Autoren bevorzugen sie es, alleine die Nutzung zu erlernen und ziehen Handbücher dem Versuch-und-Irrtum-Ansatz vor. Handbücher wurden meist präferiert, da diese in der Regel für bestimmte Aufgaben Schritt-für-Schritt-Anleitungen für ein spezifisches Gerät beinhalten.

In einer Studie von Jones et al. (2013) stellte sich heraus, dass Tablet-PCs (iPads) das Engagement und die Motivation der Senioren zwar erhöhten, allerdings das Gewicht der Geräte für ältere Menschen problematisch sein kann, wenn sie diese in den Händen halten müssen. Generell ist jedoch direkte Interaktion mittels Berührung, die von Touchscreens ermöglicht wird, von Vorteil, da Senioren laut Motti et al. (2013) weniger gehemmt sind mit einem entsprechenden System zu interagieren. Die Autoren stellten allerdings bei der Touch-Eingabe auch einige Probleme, wie z. B. beim Timing von Tipp-Gesten, fest.

Privatsphäre

Morris et al. (2013) untersuchten eine große Anzahl an Publikationen zum Thema Smarthome-Technologie zur Verbesserung der Gesundheit und des Wohlbefindens von älteren Menschen. Bei der Akzeptanz der Technologie spielten laut den Autoren die Wahrung der Privatsphäre sowie Sicherheitsbedenken eine wichtige Rolle. Diese Erkenntnis ist übereinstimmend mit den bereits vorgestellten Ergebnissen der Seniorenbefragung.

9.1.3 Ergebnisse

Die speziellen Anforderungen der Zielgruppe Senioren wurden gesammelt, ausgewertet und bei der Entwicklung des System-Konzepts berücksichtigt. Bei der Anforderungsanalyse wurde festgestellt, dass als Interaktionsgerät digitale Bilderrahmen gut geeignet sein können, da diese den Senioren häufig bereits bekannt sind und sich problemlos in eine bestehende Wohnung integrieren lassen. Zudem wurden in der Literatur bereits Systeme mit Bilderrahmen für Senioren vorgestellt. Diese Systeme konnten zum Teil durchaus einen Beitrag zur Verbesserung des menschlichen Wohlbefindens leisten. Eine Wahrung der Privatsphäre war Senioren wichtig, sodass diese besonders berücksichtigt werden musste.

Auf dieser Basis wurde das Konzept für einen Hardware- und Software-Prototyp und der Interaktion mit diesem entwickelt, der nachfolgend als CARE-Prototyp oder CARE-System bezeichnet wird. Im Wesentlichen bestand das System aus Sensorik, Logik zur Interpretation des Kontexts und der Empfehlungsauswahl sowie Bildschirmausgabe.

9.2 Empfehlungen zur Verbesserung des Wohlbefindens

9.2.1 Modell des Wohlbefindens für CARE

Für die Generierung und Auswahl von Empfehlungen wurde für den CARE-Prototyp das Wohlbefindensmodell herangezogen. Die Entwicklung des Modells wurde bereits in Kapitel 2 behandelt. Abbildung 9.4 zeigt die verschiedenen Kategorien des Wohlbefindens mit Unterkategorien, die direkt für das CARE-Empfehlungssystem umgesetzt wurden.

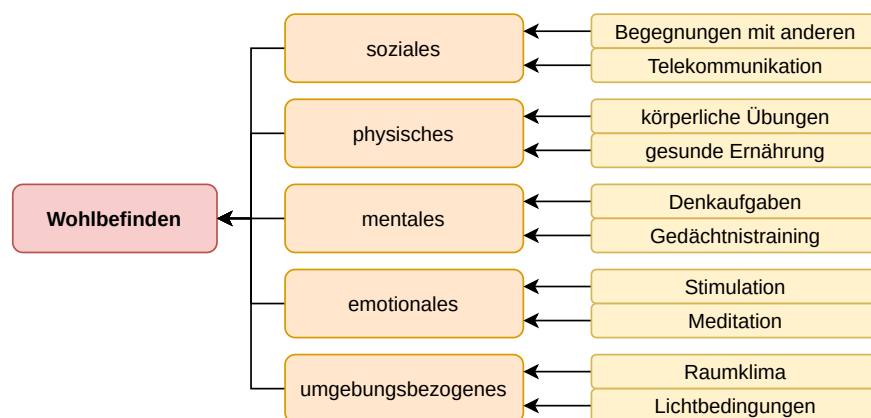


ABBILDUNG 9.4: Das für die Empfehlungen bei CARE verwendete Wohlbefindensmodell.

Um die Plausibilität des Wohlbefindensmodells von CARE zu evaluieren, wurden strukturierte Interviews mit einer Peergruppe aus 21 Senioren (13 Frauen, 8 Männer) im Alter zwischen 67 und 83 Jahren durchgeführt. Alle Teilnehmer lebten entweder alleine oder mit einem Partner in einer Wohnung oder einem Haus. Die Ergebnisse der Befragungen, die in der eigenen Arbeit (Rist, Seiderer & André, 2018) veröffentlicht wurden, werden im Folgenden zusammengefasst.

Die Antworten der Studienteilnehmer auf Fragen bezüglich der physischen Fitness zeigten, dass Personen die angaben, häufiger körperliche Aktivitäten durchzuführen, sich auch fitter als andere Senioren fühlten. Die meisten Teilnehmer wollten mehr physische Aktivitäten betreiben und stärker auf ihre Ernährung achten. Körperliche Einschränkungen wurden als häufigster Grund erwähnt, warum zu wenige Aktivitäten durchgeführt werden. Da die Senioren für ein gutes Wohlbefinden gesundes Essen explizit erwähnten, wurde dieses als Teil des **physischen Wohlbefindens** im Modell aufgenommen. Bei den Fragen zum **mental**en Wohlbefinden merkten die meisten Senioren an, dass sie gezielt Aktivitäten betreiben, um ihre mentale Fitness zu verbessern. Hierzu zählten beispielsweise Rätselaufgaben (Quiz), Kreuzworträtsel und das Erinnern an Nummern. Bezüglich des **sozialen Wohlbefindens** wurden z. B. Fragen gestellt, wie oft andere Personen getroffen werden oder mit diesen telefoniert wird. Für die Teilnehmer war der Kontakt zu Familienmitgliedern und Freunden sehr wichtig. Als Barrieren wurden hierbei fehlende Transportmittel und Müdigkeit genannt.

9.2.2 Umsetzung

Die Grundannahme für die Umsetzung bestand darin, dass geeignete Empfehlungen sich positiv auf das Wohlbefinden von Senioren auswirken, falls sie für die Person aktuell relevant sind, die Empfehlungen vom Nutzer ausführbar sind und Motivation besteht, sie durchzuführen. Im Folgenden werden die Gestaltung und der Inhalt der Empfehlungen der fünf Wohlbefindenskategorien behandelt.

Gestaltung Abbildung 9.5 links zeigt eine Empfehlung des CARE-Systems (Prototyp 2). Sie gehörte in diesem Fall zur Kategorie „physisches Wohlbefinden“. Generell waren Empfehlungen so aufgebaut, dass sie oben links ein Symbol für die entsprechende Kategorie und eine einheitliche Farbcodierung für den Hintergrund und das Symbol verwendeten. Jede Empfehlung besaß zudem einen Titel und einen Text, der erklärte, was der Nutzer tun sollte. Ein Bild oder eine einfache Animation veranschaulichte die Tätigkeit. Spätere Versionen des CARE-Systems konnten zusätzlich interaktive Elemente enthalten.



ABBILDUNG 9.5: Links: Beispiel für eine körperliche Übung. Rechts: Beispiel für eine Empfehlung der Kategorie „Ernährung“, die dem physischen Wohlbefinden untergeordnet war.

Physisches Wohlbefinden Physisches Wohlbefinden bestand beim verwendeten Wohlbefindensmodell aus den Unterkategorien „physische Fitness“ und „Ernährung“.

Die physische Fitness, im Sinne von Ausdauer, Muskelstärke, Koordination und Beweglichkeit kann durch physische Aktivitäten verbessert werden. Für Senioren existieren hinsichtlich geeigneter Übungen eine Vielzahl an Quellen in Form von Büchern, Zeitschriften, TV-Beiträgen und

Websites. Bei CARE waren Empfehlungen enthalten, die motivieren sollten, körperliche Aktivitäten durchzuführen. Zusätzlich wurden spezielle Übungen in Empfehlungsform vorgeschlagen. Zu den physischen Aktivitäten zählten beispielsweise das Spaziergehen im Park. Die Übungen dienen einem bestimmten Ziel z. B. der Verbesserung der Fingerbeweglichkeit (siehe Abbildung 9.5 links).

Gemäß dem Modell waren in dieser Kategorie des Wohlbefindens ebenfalls Empfehlungen für gesunde Ernährung enthalten (siehe Abbildung 9.5 rechts). Hierzu zählte beispielsweise regelmäßiges Trinken, da Senioren laut Begum und Johnson (2010) durch veränderte Körperfunktionen leichter zu Dehydrierung neigen und zudem ein verringertes Durstgefühl besitzen.



ABBILDUNG 9.6: Links: Beispiel für ein personalisiertes, interaktives Quiz von Prototyp 2, zur Verbesserung der Erinnerungsfähigkeit. Rechts: Empfehlung mit expliziter Aufforderung zum Lachen.

Mentales Wohlbefinden Physische Aktivität besitzt auch positive Auswirkungen auf den geistigen Zustand. In einer Studie von Larson et al. (2006) mit 1740 Senioren (älter als 65 Jahre) wurde festgestellt, dass Demenz bei älteren Menschen, die mindestens dreimal in der Woche einen Trainingslauf absolvierten, im Vergleich zu Studienteilnehmern, die weniger oft liefen, 35 % seltener auftrat.

Übungen zur Verbesserung des mentalen Wohlbefindens waren inspiriert vom „Mini-Mental-Status-Test“ (MMST) (engl. Mini Mental State Examination), der als Hilfsmittel zur Bestimmung des mentalen Zustands eines Seniors von Pflegern eingesetzt werden kann (Folstein et al., 1983). Der MMST besteht aus 11 Aufgaben, die folgende fünf Kategorien abdecken: Orientierung, Merkfähigkeit, Aufmerksamkeit und Rechenfähigkeit, Erinnerungsfähigkeit und Sprache. In CARE wurden beispielsweise Rätselaufgaben (Quiz) integriert und soweit wie möglich für den Nutzer personalisiert. Damit sollten sie für den Anwender relevanter und attraktiver sein (siehe Abbildung 9.6 links).

Emotionales Wohlbefinden Für eine Verbesserung des emotionalen Wohlbefindens wurden Empfehlungen so gestaltet, dass sie möglichst einen positiven Affekt auslösen oder negative Gefühle reduzieren. Beispielsweise besitzt Lachen einen vorteilhaften Effekt auf die Gesundheit und kann negative Gefühle reduzieren (Yim, 2016). Eine diesbezügliche therapeutische Nutzung ist ebenfalls möglich. Von Tsujita und Rekimoto (2011) wurden zum Beispiel Systeme vorgestellt, die den Nutzer dazu bewegen sollten, gezielt zu lachen. Gemäß diesem Prinzip wurde eine Empfehlung (siehe Abbildung 9.6 rechts), die dem Nutzer explizit empfiehlt zu lachen, integriert. Es waren zusätzlich mehrere Witze enthalten, die spontanes Lachen auslösen konnten.

Soziales Wohlbefinden Um das soziale Wohlbefinden zu verbessern, waren in CARE Empfehlungen enthalten, die das Potenzial besaßen, die Beziehung zu anderen Personen positiv zu beeinflussen. Beispielsweise konnte das System vorschlagen, Aktivitäten mit Familienmitgliedern

oder Freunden, wie einen Spaziergang oder Ausflug, zu planen. Abbildung 9.7 links zeigt hierzu ein Beispiel.

Umgebungsbezogenes Wohlbefinden Empfehlungen dieser Kategorie bezogen sich auf das Umfeld des Nutzers, wozu auch die eigene Wohnung zählte. Beispielsweise waren Empfehlungen zur Verbesserung des Raumklimas durch regelmäßiges Lüften enthalten (siehe Abbildung 9.7 rechts).



ABBILDUNG 9.7: Links: Vorschlag, mit Freunden einen Ausflug in die Stadt zu unternehmen. Rechts: Empfehlung den Raum regelmäßig zu lüften.

9.3 Sensorik für das CARE-System

Um kontextsensitive Empfehlungen ausgeben zu können, ist es notwendig, Kontextinformationen mithilfe von Hardware und virtuellen Sensordaten zu ermitteln. Hierfür wurden für das CARE-System geeignete Sensoren unter folgenden Gesichtspunkten ausgewählt:

- Die Sensorik sollte von körperlich relativ fitten Senioren problemlos nutzbar sein. Sie sollte von diesen getragen oder möglichst unaufdringlich in deren Wohnung integriert werden können.
- Aus Gründen des Schutzes der Privatsphäre war es unerwünscht, permanent Kamera- oder Audioaufnahmen zu erstellen.
- Sensordaten sollten von einem Tablet mit vergleichsweise geringer Leistung auswertbar sein.
- Die Wohnung sollte nur temporär und möglichst geringfügig umgestaltet werden müssen.
- Die Sensorik sollte mit einer Anleitung unkompliziert zu installieren sein und auch bei einem Transport nicht beschädigt werden.
- Das System sollte auch ohne Internetverbindung lauffähig sein, da derzeit, ältere Senioren häufig über keinen Internetzugang verfügen³.
- Die Kosten des Gesamtsystems sollten nicht zu hoch sein (kleiner ca. 2000 €).

9.3.1 Stationäre Sensorik

Raumklima, Raumpräsenz, Nutzung elektrischer Geräte

Um generelle Informationen über die Präsenz von Personen im Raum, die Lufttemperatur / Luftfeuchtigkeit, den Öffnungszustand der Fenster und Nutzung bestimmter elektrischer Geräte zu

³laut D21-Digital-Index waren in Deutschland 2020 ca. 52 % der über 70-Jährigen online; von der Gesamtbevölkerung waren es ca. 88 %: https://initiated21.de/app/uploads/2021/02/d21-digital-index-2020_2021.pdf (08.04.2021)

erhalten, kann kommerzielle Sensorik von verschiedenen Smarthome-Systemen eingesetzt werden. Viele der Systeme können auch offline betrieben werden, insbesondere falls alternative Steuerungssoftware eingesetzt wird. Für das CARE-System wurde das Homematic-System verwendet, das von der quelloffenen Software FHEM unterstützt wird, die unter Linux und Windows lauffähig ist. Als Hardware wird lediglich eine Bridge oder ein spezieller USB-Stick benötigt, um mit den drahtlosen Komponenten zu kommunizieren. Hierbei ist einmalig ein „Pairing“ (Kopplung) mit den Sensoren nötig. Die Kopplung sowie evtl. Einstellungen an den Sensoren bleiben jedoch dauerhaft erhalten. Ein Austausch der Batterien ist je nach Nutzung, Einstellungen und Sensortyp meistens erst nach mehreren Jahren erforderlich. Die funkgesteuerten, drahtlosen Sensoren und Aktuatoren können relativ problemlos in einer Wohnung installiert werden und zeigten sich in eigenen Tests durch hohe Zuverlässigkeit aus. Aus diesen Gründen wurden sie für die Verwendung bei der Entwicklung der CARE-Prototypen ausgewählt.

Zur Überwachung der Raumluft sind bereits weniger komplexe Gassensoren ausreichend, die auch immer häufiger bei Smarthome-Systemen eingesetzt werden. Es wurden beim CARE-System kostengünstige VOCs-Sensoren eingesetzt, die viele Arten von Gerüchen sowie Kohlendioxid messtechnisch erfassen können. Die Genauigkeit der Messergebnisse ist mittelmäßig, jedoch war sie für die Verwendung bei den CARE-Prototypen ausreichend. Zudem konnten sie bei hinreichender Luftzirkulation in die Prototypen verbaut werden. Ein unkomplizierter Anschluss per USB ist, je nach Modell, möglich.

Systeminteraktion

Als Basis der CARE-Prototypen wurden Tablet-PCs eingesetzt, da diese kompakt, stromsparend sowie relativ preisgünstig sind und bereits einen Touchscreen zur Interaktion bieten. Zusätzlich verfügen diese meistens über eine Kamera, Mikrofon und weitere Sensorik. Aus Gründen des Schutzes der Privatsphäre war die Nutzung der Kamera oder des Mikrofons nur eingeschränkt möglich gewesen.

Die explizite Interaktion mit dem System sollte erst dann beginnen, wenn sich eine Person vor dem Tablet-PC befand (Nutzerpräsenz). Für diese Erkennung wurden IR- bzw. Ultraschall-Abstandssensoren eingesetzt, die relativ einfach in den Prototyp integrierbar waren. Diese können allerdings nur einen begrenzten räumlichen Bereich überwachen. Im Gegensatz zu PIR-Sensoren besitzen sie jedoch den Vorteil, dass sie nicht ausschließlich auf Bewegungen reagieren, die beim Stehen vor dem System nicht ausreichend vorhanden wären. Eine Alternative wären grundsätzlich auch Radarsensoren gewesen, die auch auf kleinere Bewegungen reagieren, allerdings müsste hierbei der Erkennungsbereich sauber definiert werden, da diese auch Bewegungen hinter Objekten oder Wänden erkennen, was in diesem Fall unerwünscht war. Auch IR-Abstandssensoren besitzen Nachteile, sie können beispielsweise fehlerhafte Abstandswerte liefern, falls sie gegenüber einer Fensterscheibe positioniert werden (Reflexion), Sonnenlicht (enthält IR-Strahlung) auf sie einfällt oder ein sehr dunkler Gegenstand, wie ein Kleidungsstück, sich vor ihnen befindet. Zusätzlich können viele Sensor-Modelle sehr geringe Abstände, die bei der Interaktion mit dem System auftreten können, nicht erfassen. Ultraschall-Abstandssensoren können den Abstand von Objekten, die den Schall stark ablenken schlecht erkennen oder werden durch Schallreflexionen gestört. Sehr nahe Objekte stellen auch hier ein Problem dar. Für CARE-Prototyp 1 wurde primär ein IR-Abstandssensor für die Nutzerpräsenz genutzt. Dieser zeigte zwar die bereits erwähnten Probleme, wobei diese jedoch teilweise mit Logik bzw. dem Aufstellungsort reduziert werden konnten. Bei Prototyp 2 wurde eine Kombination aus IR- und Ultraschall-Abstandssensoren eingesetzt, die eine robustere Erfassung ermöglichte.

Überwachung von Übungen

Für die Erkennung, ob eine Person vor dem CARE-Prototyp steht (Nutzerpräsenz), könnten grundsätzlich auch Druckmatten genutzt werden, die prinzipiell auch für die Überwachung

physischer Übungen geeignet sind. Bei den CARE-Prototypen (1 bis 2) wurden sie jedoch für die Feldstudien nicht eingesetzt, da sie konstruktionsbedingt leicht beschädigt werden können und ihre Verkabelung die Gefahr birgt, dass Personen darüber stolpern. Bei den Untersuchungen mit dem CARE-Prototyp 2 mit Gamifizierung wurde allerdings eine sogenannte „Tanzmatte“, die robust und weniger komplex aufgebaut ist, verwendet.

Eine Überwachung mittels stationärer Sensorik war bei den angedachten physischen Übungen ohne die Nutzung von Kameras kaum robust möglich und zudem nicht das Hauptziel des CARE-Systems. Alternativ wurden tragbare Sensorgeräte untersucht, die zum Teil eine Überwachung einzelner Übungen und das Erfassen von Informationen über die körperliche Aktivität der Senioren ermöglicht hätten.

9.3.2 Mobile Sensorik

Im Rahmen des CARE-Systems waren grundsätzlich tragbare Geräte für die Überwachung einiger physischen Übungen geeignet. Bei vielen tragbaren kommerziellen Geräten ist es jedoch nicht möglich, die benötigten Sensordaten mittels eigener Software in Echtzeit auszuwerten, was eine Integration in ein eigenes System deutlich erschwert oder verhindert. Bei den nachfolgend aufgeführten Geräten war ein Auslesen der Daten mit eigener Software, in Echtzeit (bis auf den AX3-Sensor), möglich.

Für einen ersten Test wurden Sensoren von Axivity eingesetzt, die mit eigener Software ausgelesen werden konnten. Der kompakte **AX3-Sensor**⁴ (23×32,5×7,6 mm) konnte mit einem Armband verwendet werden. In Abbildung 9.8 links ist der aus dem Armband herausnehmbare Sensor dargestellt. Der AX3-Sensor enthielt ein 3-Achsen-Accelerometer, einen Helligkeits- und Temperatursensor. Die Accelerometer-Daten konnten laut Hersteller-Website bei einer Abtastrate von 100 Hz bis zu 14 Tage aufgezeichnet werden. Die Daten konnten bei diesem Gerät ausschließlich per USB ausgelesen werden. Zudem existierte eine Version mit Bluetooth LE (**WAX9**⁵), die die Daten in Echtzeit an ein anderes Gerät senden konnte. Hierbei waren folgende Sensoren vorhanden: 3-Achsen Accelerometer / Gyroskop / Magnetometer, Temperatur- und Luftdrucksensor.



ABBILDUNG 9.8: Links: Axivity WAX-9-Sensor (AX3-Sensor äußerlich baugleich). Rechts: Microsoft Band 2.

Im Rahmen der Evaluation des CARE-Prototyp 1 wurden den beiden Studienteilnehmern (76 und 75 Jahre alt) jeweils ein geladener AX3-Sensor mit Armband gegeben und die Senioren gebeten, diese möglichst lange zu tragen. Für einen ersten Test waren die AX3-Sensoren gut geeignet, da diese eine lange Laufzeit besaßen und keine weitere Interaktion während der Studiendauer, wie beispielsweise Aufladen, nötig war. Eine Befragung und Datenanalyse wurde nach der Studie durchgeführt. Die Befragung zeigte, dass die tragbaren AX3-Sensoren von den Senioren grundsätzlich positiv aufgenommen wurden. Allerdings stellte sich bei der Befragung sowie der Sensordatenauswertung heraus, dass die Senioren häufig vergessen hatten, die Sensoren zu tragen.

Für eine Integration in das CARE-System zur Echtzeitüberwachung von physischen Übungen wäre die Bluetooth-Version (WAX9) des AX3-Sensors prinzipiell geeignet gewesen, für die auch ein Plugin für das Echtzeitsignalverarbeitungs-Framework SSI entwickelt wurde (siehe Anhang C).

⁴<https://axivity.com/product/ax3> (03.12.2019)

⁵<https://axivity.com/downloads/wax9> (03.12.2019)

Von einer diesbezüglichen Integration wurde dennoch abgesehen, da die Senioren bereits die AX3-Sensoren vergessen hatten zu tragen und mit den WAX9-Sensoren in Tests keine automatische, robuste Bluetooth-Verbindung und Datenübertragung mit einem Windows-System hergestellt werden konnte, wenn diese am Handgelenk getragen wurden. Im Vergleich zu den AX3-Sensoren wäre zudem nach einigen Stunden Nutzungsdauer ein Laden der Akkus der WAX9-Sensoren notwendig gewesen.

Parallel wurde ebenfalls die Nutzung eines **Myo-Armbandes** von Thalmic Labs⁶ (wird laut Website seit Oktober 2018 nicht mehr hergestellt) untersucht. Dieses konnte durch die Auswertung der Muskelaktivität (8-Kanal EMG-Sensor) und Bewegungen (3-Achsen Accelerometer / Gyroskop / Magnetometer) bestimmte Finger und Arm- / Hand-Gesten erkennen. Die Daten wurden hierbei per Bluetooth übertragen. Durch die Nutzung der Elektromyografie-Daten wäre dies grundsätzlich eines der wenigen Sensorgeräte gewesen, das es erlaubt hätte, Fingergesten ohne Kameradaten zu erfassen. Dies hätte eine Überprüfung der Fingerübungen beim CARE-System ermöglicht. Bei diesem innovativen Gerät zeigten sich allerdings eine Vielzahl an praktischen Hindernissen, die eine Nutzung beim CARE-Prototyp verhinderten. Generell war dieses Gerät, im Vergleich zum WAX9-Sensor oder einem Fitness-Tracker, für ein dauerhaftes Tragen nicht vorgesehen, da es im Alltag störte. Bereits das Anlegen des Armbandes hätte für Senioren ein Hindernis darstellen können und benötigte ein kurzärmeliges Oberteil oder eines mit weiten Ärmeln, da das Gerät direkt auf der Haut des Unterarms fest, in einer bestimmten Orientierung, platziert werden musste. Längeres Tragen war bei eigenen Versuchen zudem unangenehm, da der Unterarm taub werden konnte. Zusätzlich war es für viele Gesten auf Basis der EMG-Daten notwendig, nach jedem Anziehen des Gerätes eine Kalibrierung durchzuführen, da die Daten stark vom Nutzer und der exakten Sensorposition abhingen. Auch die vom Hersteller bereitgestellten allgemeineren Modelle funktionierten hierbei häufig nur begrenzt. Zudem war auch bei diesem Gerät der Aufbau einer Bluetooth-Verbindung regelmäßig mit höherem Nutzaufwand verbunden und es musste in kürzeren Abständen aufgeladen werden. Daher wäre eine eigenständige Nutzung durch Senioren nicht sinnvoll möglich gewesen.

Einer der wenigen kompakten Fitness-Tracker, dessen Daten in Echtzeit auf einem Smartphone per Bluetooth auslesbar waren, war das **Microsoft Band 2** (siehe Abbildung 9.8 rechts), das allerdings erst im November 2015 in Europa⁷ (nicht offiziell in Deutschland) auf den Markt kam. Es konnte daher erst nach Ablauf des CARE-Projektes (Dezember 2015), bei den weiterführenden Untersuchungen berücksichtigt werden. Die Produktion wurde bereits im Oktober 2016⁸, ohne Nachfolger, offiziell eingestellt. Das Microsoft Band 2 war hinsichtlich der verbauten Sensorik von Interesse, da es eines der wenigen Geräte war, die Hautleitwert-Daten (EDA) liefern konnten, was beispielsweise zur Überwachung des Stresslevels bei einer Aktivität nützlich sein kann. Im Rahmen der weiterführenden Untersuchungen wurden jedoch ausschließlich die 3-Achsen-Accelerometer-Daten des Gerätes für Gestenerkennung verwendet, um eine Übung zu überwachen, die mit Senioren in einem Workshop evaluiert wurde. Hierfür war von Bedeutung, dass ein möglichst kompaktes, tragbares, robustes Consumer-Gerät genutzt wurde, um die Akzeptanz im Rahmen von Gamifizierungen für das CARE-System untersuchen zu können. Die Ergebnisse werden in einem späteren Abschnitt beschrieben.

9.4 Entwicklung CARE-Prototyp 1

In diesem Abschnitt wird die Implementierung und Evaluation des ersten CARE-Prototyps beschrieben. Teile hiervon wurden in den eigenen Publikationen (Rist, [Seiderer](#), Hammer, Mayr

⁶<https://support.getmyo.com/hc/en-us> (03.12.2019)

⁷<https://www.techadvisor.co.uk/news/wearable-tech/microsoft-band-2-release-date-price-specs-18nov-3626324/> (03.12.2019)

⁸<https://www.theverge.com/circuitbreaker/2016/10/3/13152590/microsoft-band-discontinued> (03.12.2019)

& André, 2015; Seiderer, Hammer, André, Mayr & Rist, 2015) veröffentlicht. Die später entwickelten Prototypen basierten auf diesem, wurden jedoch durch Erkenntnisse, die während der durchgeführten Langzeit-Feldstudie entstanden, weiterentwickelt.

9.4.1 Umsetzung

Hardware

Für den ersten Prototyp wurde als Tablet ein preisgünstiges (ca. 400 €) HP Omni 10 5600eg (CPU: Intel Atom Z3770, RAM: 2 GiB, Speicher: 32 GiB) mit Windows 8.1 verwendet, das in eine magnetische Pinnwand integriert wurde. In Abbildung 9.9 ist der Aufbau des CARE-Prototyps 1 dargestellt. Abbildung 9.10 zeigt die Vorderseite der selbst entwickelten Sensorplatine. Ein Überblick über das System ist in Abbildung 9.11 dargestellt. Die Pinnwand bestand aus einem Metallgitter, das mit einer Filzfolie beklebt wurde. Die Pinnwand konnte mithilfe von zwei Regalschienen an einer Wand befestigt und in der Höhe verstellt werden. Die komplette Sensorik war auf der Rückseite der Pinnwand bzw. im Tablet enthalten. Das Tablet verfügte über ein Accelerometer, Gyroskop, Inklinometer, Mikrofon, einen Kompass, Lichtsensor und eine Frontkamera. Mittels eines USB-Hubs waren zusätzlich ein AMS iAM-VOCs-Gassensor sowie ein Arduino Nano mit analogem IR-Abstandssensor (Sharp GP2Y0A02YK0F) mit einem Erkennungsbereich von 20 bis 150 cm, ein Raumtemperatur- / Raumluftfeuchtigkeit-Sensor (DHT22) sowie ein 3-Achsen-Magnetometer (Honeywell HMC5883L⁹ (Platine: GY-271)) mit dem Tablet verbunden. Das Magnetometer wurde auf der Rückseite des Prototyps im unteren Bereich platziert. Auf der Vorderseite hätten Systemnutzer diesen Bereich mit Magneten als Pinnwand verwenden können. Das Magnetometer ermöglichte somit eine grundsätzliche Überwachung von Interaktionen mit der Pinnwand, da es auch kleine Änderungen im Magnetfeld erfassen konnte.

Der Prototyp konnte in einer Wohnung sehr schnell aufgebaut werden, da ausschließlich eine naheliegende Steckdose sowie eine freie Wandfläche mit einer Breite von ca. 35 cm benötigt wurden. Externe Sensorik war bei diesem ersten Prototyp nicht vorhanden.

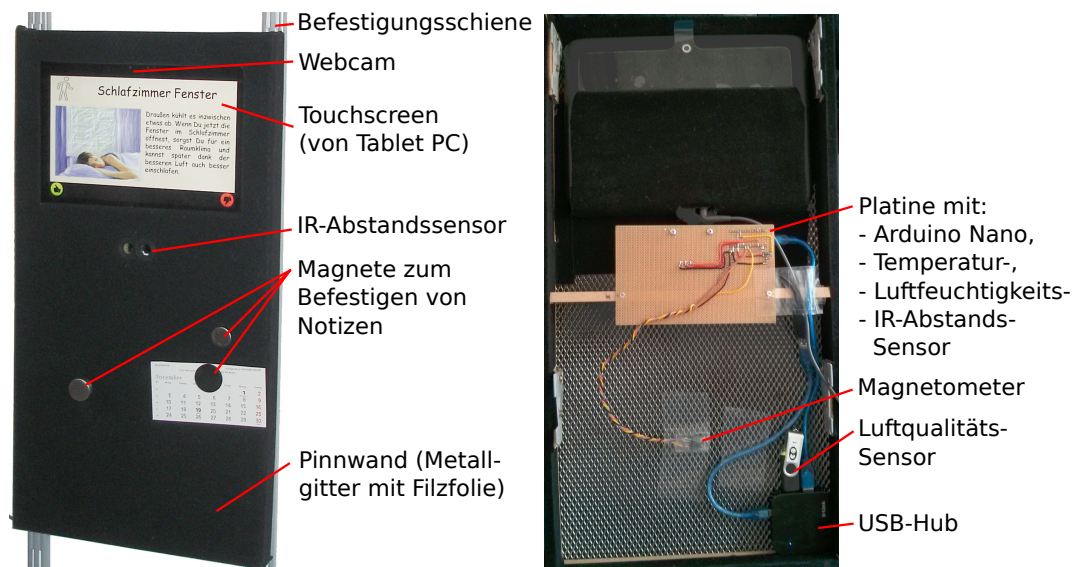


ABBILDUNG 9.9: Aufbau von CARE-Prototyp 1.

⁹https://cdn-shop.adafruit.com/datasheets/HMC5883L_3-Axis_Digital_Compass_IC.pdf (13.03.2021)

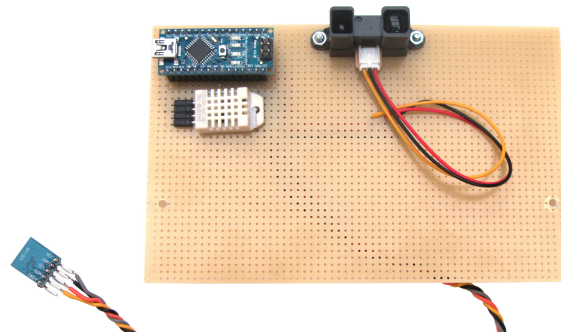


ABBILDUNG 9.10: Vorderseite der Platine mit Arduino Nano und Sensorik.

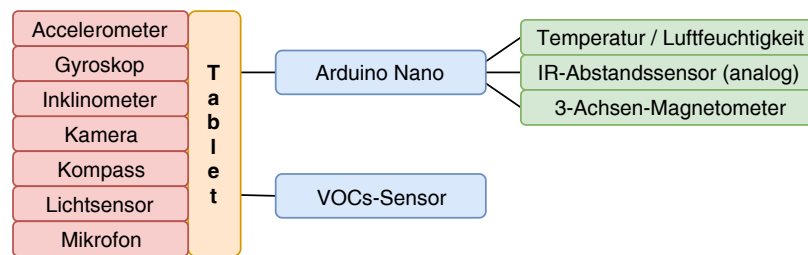


ABBILDUNG 9.11: Hardware-Aufbau CARE-Prototyp 1.

Software

In diesem Abschnitt werden die Softwarekomponenten des CARE-Systems vorgestellt. Einige hiervon wurden in veränderter, angepasster Form auch für den zweiten Prototyp eingesetzt. Die Grundstruktur war daher größtenteils identisch und es werden bei der Beschreibung des zweiten Prototyps primär Änderungen aufgeführt.

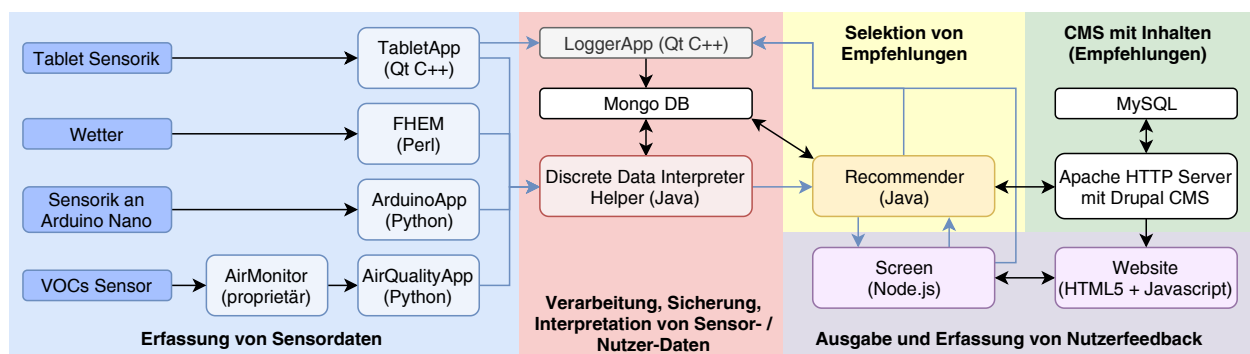


ABBILDUNG 9.12: Übersicht der Programme des CARE-Prototyps 1, die auf dem Tablet ausgeführt wurden.

Übersicht In Abbildung 9.12 sind im Detail die einzelnen Programme und deren interne Kommunikation mit ihren Programmnamen und der Programmiersprache dargestellt. Bis auf die Softwares „FHEM“, „AirMonitor“, „Mongo DB“, „MySQL“ und „Apache“, wurden die Programme für den CARE-Prototyp neu entwickelt. Die gewählten Programmiersprachen hingen hierbei von verfügbaren Bibliotheken und dem Programmumfang ab. Der generelle Aufbau der Softwarekomponenten kann in die Bereiche „Datenerfassung“ (in der Abbildung blau), „Datenverarbeitung, Sicherung und Interpretation“ (rot), „Empfehlungsauswahl“ (gelb), „Datenbank für Inhalte (CMS)“ (grün) und „GUI“ (violett) unterteilt werden.

Um die Handhabung des Systems zu erleichtern, wurden alle Programme von einem Startskript in einer festgelegten Reihenfolge, teilweise mit Verzögerung, gestartet. Das Startskript konnte zudem notfalls abgestürzte Programme neu starten. Zusätzlich war es optional möglich, Sensordaten in einem Webinterface zu visualisieren, um Probleme feststellen zu können.

Blaue Verbindungslinien kennzeichnen die Kommunikation mittels des per TLS verschlüsselten XMPP-Protokolls, das auch die Übertragung kleinerer Binärdateien, wie komprimierten Fotos, erlaubt. Es wurde in den meisten Fällen die Publish / Subscribe-Funktionalität von XMPP verwendet, wobei die Nachrichten im JSON-Format kodiert waren. Als XMPP-Server wurde das quelloffene „Openfire“¹⁰ auf dem System gestartet. Externe Server wären allerdings ebenfalls nutzbar gewesen, falls Daten von einem anderen Rechner sicher über das Internet hätten übertragen werden müssen. Insbesondere die Komponenten der Sensordatenerfassung mussten flexibel gehandhabt werden, da nicht immer eine Internetverbindung vorhanden war, wodurch die Wetterdaten durch FHEM nicht abgerufen werden konnten. Das System musste daher auch mit fehlenden Informationen lauffähig bleiben. Zusätzlich sollte die Option bestehen, weitere Sensorquellen zu ergänzen. Insbesondere beim ersten Prototyp musste darauf geachtet werden, dass das System erweiterbar und anpassbar war, da sich die Anforderungen nach der Nutzerevaluation stärker verändern konnten. Daher war hier die Anzahl der einzelnen Programme höher als in späteren CARE-Prototypen. Generell waren sehr viele Funktionen des Systems durch Konfigurationsdateien anpassbar, um diese leichter testen zu können.

Erfassung, Aufbereitung, Interpretation und Speicherung von Sensordaten In Abbildung 9.12 auf der linken Seite sind die Programme, die Sensordaten abriefen blau hinterlegt. Wetterdaten wurden als virtueller Sensor behandelt (erfasst von FHEM). Zusätzlich wurde die Tageszeit mittels des berechneten Sonnenaufgangs und Untergangs an den aktuellen Koordinaten bestimmt. Als virtueller Sensor können auch die Nutzereingaben gezählt werden, die allerdings über die Ausgabekomponenten (*Website* und *Screen*) sowie das Programm zur Empfehlungsauswahl (*Recommender*) in die Datenbank mit Sensordaten geleitet wurden.

Sensoren, die mit dem Arduino Nano verbunden waren, wurden von der *ArduinoApp* mit 2 Hz abgetastet. Der iAM VOCs-Sensor wurde mit 1 Hz von der proprietären *AirMonitor*-Software ausgelesen, die *AirQualityApp* sendete die Daten per XMPP weiter. Tablet-Sensorik, die höhere Abtastraten als 1 Hz besaß, wurde durch die *TabletApp* mittels Berechnung von Mittelwerten auf 1 Hz reduziert (Downsampling). Dies galt beispielsweise auch für die Audiodaten, die zudem aus Gründen des Schutzes der Privatsphäre nur vorverarbeitet gesichert wurden. Es wurden hierbei ausschließlich Maximalwerte und RMS-Energiewerte berechnet und weiterverarbeitet. Optional konnte die *TabletApp* auch Bilder von der Videokamera erfassen. Diese wurden allerdings ausschließlich aufgenommen, wenn das System eine Nutzerpräsenz vor dem System erkannt hatte. Es wurden daraufhin im 10-Sekunden-Intervall mit einer Maximalanzahl von 50 Bildern pro Sitzung, Bilder in VGA-Auflösung als JPEG-Dateien, aufgenommen. Diese Bilder dienten der manuellen Auswertung und Überprüfung des Systems und wurden nur aufgezeichnet, falls die Studienteilnehmer zuvor zusätzlich ausdrücklich zugestimmt hatten. Die Bilder wurden an die *LoggerApp* gesendet, die Log-Daten (z. B. Programmausgaben) verschiedener Programme gesondert in der MongoDB-Datenbank mit Zeitstempeln zur späteren Auswertung speicherte.

Alle Sensordaten mussten an das *Discrete Data Interpreter Helper* (DDIH) Programm per XMPP-Protokoll gesendet werden, das diese in die Datenbank „sensordataDB“ einfügte und serverseitige Javascript-Funktionen zur Berechnung von Kontextinformationen aufrief. Jede Nachricht mit Sensordaten verfügte über die JSON-Struktur, die an einem Beispiel in Abbildung 9.13 dargestellt ist und so auch in der Datenbank abgebildet wurde. Es sind zunächst der Zeitstempel, der Wert und optional eine Zuverlässigkeit angegeben. Der Sensorwert konnte verschiedene Datentypen annehmen. Die Zuverlässigkeit wäre unter anderem bei Klassifikationsergebnissen von Relevanz

¹⁰<https://www.igniterealtime.org/projects/openfire/> (06.12.2019)

gewesen. Zusätzlich waren Informationen über den Raum, in dem der Sensor platziert war, enthalten, da die Berechnungen der Kontextinformationen, wie dem Raumklima, für mehrere Räume möglich gewesen wäre. Es konnte der Raumtyp angegeben werden, falls bei der Berechnung nur Sensoren aus dem Wohnzimmer verwendet werden sollten sowie eine ID, da es z. B. mehrere Schlafzimmer in einer Wohnung geben könnte. Schließlich waren Daten über den Sensor in jeder Nachricht vorhanden. Diese beinhalteten den Sensortyp, den Positionstyp inklusive einer ID sowie dem Namen der Software von der die Daten stammten. Diese Informationen waren beispielsweise relevant, um unabhängig von bestimmten Sensoren Informationen, wie die mittlere Raumtemperatur zu berechnen, falls mehrere Sensoren dort vorhanden sein sollten. Die Positionsinformation konnte eingesetzt werden, um die Auswahl der Sensoren weiter einzuschränken.

Die Sensordaten enthielten bezüglich des Raumes und des Sensors redundante Informationen, daher war es möglich, automatisch neue Sensoren hinzuzufügen, ohne dass diese zuvor manuell, zentral angelegt werden mussten. Beim Abspeichern in die Datenbank wurden die zusätzlichen Daten entfernt und stattdessen die IDs für den zugehörigen Sensor und Raum verwendet. Die Einträge für den Sensor und Raum wurden beim ersten Datenempfang automatisch erzeugt. Die Funktionen zur Berechnung der Kontextinformationen aus den Sensordaten wurden so gestaltet, dass diese mit fehlenden Werten arbeiten konnten. Beispielsweise wurden bei (temporär) fehlenden Wetterdaten Standardwerte genutzt. Dies war allerdings bei den Daten zur Nutzerpräsenz nur begrenzt möglich.

```

{
  "date" : "2014-11-25T10:20:42.472656Z",
  "value" : 1301.0,
  "reliability" : -1.0,
  "roomType" : "livingroom",
  "roomID" : 0,
  "sensorType" : "airQuality",
  "positionType" : "wall",
  "positionID" : 0,
  "adminSoftware" : "airQualityApp"
}

```

sensor data

room

sensor

ABBILDUNG 9.13: Beispiel-Datum des VOCs-Gassensors der *AirQualityApp* im JSON-Format. Hier war keine Zuverlässigkeit (reliability) berechenbar und wurde daher als „-1“ festgelegt. Der Sensorwert betrug 1301 ppm. Die Einheiten waren beim System festgelegt.

In der Datenbank wurden ausschließlich diskrete Daten gesichert, die teilweise bereits aus kontinuierlichen berechnet wurden, sodass die Datenmenge und der Rechenaufwand reduziert werden konnten. Dies war insbesondere für die nicht besonders leistungsfähige Tablet-Hardware mit limitiertem Speicher relevant, vor allem da das System über mindestens zwei Wochen Daten speichern können musste.

Für die Speicherung der Sensordaten wurde die quelloffene NoSQL-Datenbank „MongoDB“¹¹ eingesetzt. Es handelt sich hierbei um eine dokumentenorientierte Datenbank, wodurch im Vergleich zu relationalen Datenbanken wie PostgreSQL keine Beziehungen zwischen den Daten direkt abgebildet werden und das Datenbankschema durch die Anwendungen und nicht durch den Server festgelegt wird. Zudem können Dokumente, die im JSON-Format vorliegen, direkt eingefügt werden und unabhängig voneinander strukturiert sein. Einzelne Attribute können hierbei unterschiedliche Datentypen besitzen. MongoDB hatte u. a. den Vorteil, eine höhere Geschwindigkeit gegenüber klassischen SQL-Datenbanken zu besitzen. Diese Geschwindigkeitsvorteile gehen allerdings teilweise mit einer höheren Gefahr eines Datenverlustes bei einem Systemabsturz einher. Dadurch, dass beim ersten CARE-Prototyp ein 32-Bit-Betriebssystem verwendet werden musste, konnten keine Datenbanken über 2 GiB¹² auf dem Datenträger angelegt werden. Aus diesem

¹¹<https://www.mongodb.com/de> (09.12.2019)

¹²<https://www.mongodb.com/blog/post/32-bit-limitations> (09.12.2019)

Grund wurde eine Funktion beim *DDIH* hinzugefügt, die in der Nacht dafür sorgte, dass die Daten exportiert wurden und die Datenbank mit den Sensordaten daraufhin gelöscht wurde.

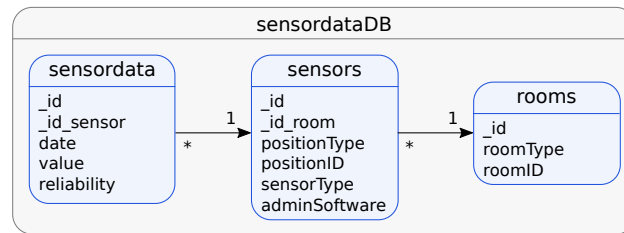


ABBILDUNG 9.14: Mongo-DB-Datenbankschema.

Die Datenstruktur der Sensordaten zeigt Abbildung 9.14. In der Datenbank „sensordataDB“ befanden sich die „Collections“ (vergleichbar mit Tabellen bei SQL-Datenbanken) „sensordata“, „sensors“ und „rooms“. Beziehungen zwischen den einzelnen Dokumenten (vergleichbar mit Zeilen bei SQL-Datenbanken) wurden hier mittels entsprechender IDs realisiert. Es war jedoch nicht direkt möglich, Abfragen über mehrere Collections in einem Befehl durchzuführen. Die flexible Handhabung der Datentypen bei Attributen (vergleichbar mit Spalten bei SQL-Datenbanken) eines Dokuments erlaubte es, dass ein Sensorwert (Attribut „value“) eine Zeichenkette oder auch numerisch sein konnte. Allerdings musste eine Funktion, die Sensordaten verarbeitete, den Datentyp kennen, um diese weiterverarbeiten zu können.

Aus den Sensordaten wurden Kontextinformationen erzeugt. Javascript-Funktionen wurden hierfür in die MongoDB-Datenbank geladen und dort vom *DDIH* aufgerufen, wenn neue Sensordaten hinzugefügt wurden. Beim ersten CARE-Prototyp wurden folgende Kontextinformationen berechnet, die bei der Empfehlungsauswahl durch den *Recommender* berücksichtigt wurden:

- Tageszeit (berechnet mittels der lokalen Uhrzeit),
- Raumklima (berechnet aus Luftqualität, Luftfeuchtigkeit und Temperatur),
- Nutzerpräsenz vor dem Demonstrator (berechnet aus den Daten des IR-Abstandssensors),
- Tageslicht (berechnet aus den geografischen Koordinaten des Standorts und der Uhrzeit) und
- Wetter (berechnet aus Daten der Yahoo-Wetter-API für den aktuellen Systemstandort und der aktuellen Uhrzeit).

Die Rückgaben der Funktion zur Bestimmung des Tageslichts waren hierbei „bright“ und „dark“. Das aktuelle Wetter wurde in fünf Stufen von sehr gut bis sehr schlecht unterteilt.

Eine Nutzerpräsenz wurde erst nach einer Erfassung von mindestens zwei Sekunden erkannt, sodass das System nicht bereits durch ein Vorbeilaufen ausgelöst wurde. Zusätzlich war eine Verzögerung integriert, um zu verhindern, dass kurzfristiges Verlassen des Sensorerfassungsbereiches oder vereinzelte fehlerhafte Messwerte direkt als Absenz eines Nutzers interpretiert wurden.

Die berechneten Kontextinformationen wurden vom *DDIH*-Programm empfangen und per XMPP an alle Subscriber verteilt. Die Daten wurden zudem in einer eigenen Collection in der Datenbank gespeichert. In diesem Fall war dies vor allem für den *Recommender* relevant.

Für das Nutzerprofil wurde der aktuelle Gemütszustand des Anwenders regelmäßig abgefragt. Es konnte zwischen „schlechter“, „neutraler“ und „guter“ Stimmung ausgewählt werden. Zusätzlich konnte jede einzelne Empfehlung optional eine positive oder negative Bewertung erhalten, die zusammen mit dem Zeitstempel und der Empfehlungs-ID abgespeichert wurden. Nutzerdaten wurden in einer gesonderten Collection („user model“) gesichert.

Empfehlungsauswahl Empfehlungen wurden bei der Erkennung einer Nutzerpräsenz durch das Programm *Recommender* in einem zweistufigen Vorgang ausgewählt. Zunächst wurden mit einem

regelbasierten Kontextfilter, auf Basis von Prolog, unpassende Empfehlungen aus der Auswahl entfernt. Wenn es beispielsweise bereits dunkel war, wurden physische Aktivitäten, wie spazieren gehen herausgefiltert. Es wurde anschließend zufällig eine Empfehlung aus der bereits gefilterten Auswahl ausgewählt. Falls eine bestimmte Empfehlung bereits dreimal in einer Stunde angezeigt wurde, wurde eine andere gewählt. Die ausgewählte Empfehlung wurde anschließend an die Anzeigekomponente weitergegeben. Die Nutzerbewertungen von Empfehlungen wurden erst in späteren CARE-Versionen berücksichtigt. Die verfügbaren Empfehlungen inklusive ihrer Bedingungen, die vom Empfehlungssystem genutzt werden, stammten aus einem Content-Management-System (CMS).

CMS und Anzeigekomponente Zur unkomplizierten, grafischen Verwaltung von Inhalten für das CARE-System wurde das Content-Management-System „Drupal“¹³ eingesetzt. Dieses benötigte einen HTTP-Server mit PHP-Unterstützung, in diesem Fall wurde Apache genutzt sowie eine MySQL- oder MariaDB-SQL-Datenbank. Unter Windows konnte hierfür das XAMPP-Paket¹⁴ portabel verwendet werden. Das CMS sollte ermöglichen, dass auch Verwandte und Bekannte der CARE-Nutzer Inhalte in das System einpflegen können. Da das CARE-System für griechische und deutsche Nutzer geeignet sein musste, waren die meisten Inhalte zweisprachig oder zusätzlich noch in Englisch vorhanden und wurden beim Start des Systems entsprechend ausgewählt. Mit dem CMS konnten Empfehlungen und die Bilder für den Bilderrahmen-Modus hinzugefügt, verändert oder gelöscht werden. Bei den Empfehlungen mussten die einzelnen Bestandteile, Auswahlbedingungen und die Wohlbefindenskategorie angegeben werden. Diese Daten wurden für die Darstellung der Empfehlungen sowie deren Filterung und Auswahl durch das Empfehlungssystem benötigt.

Das CMS wurde zusätzlich auf einem über das Internet zugreifbaren Linux-Server betrieben, sodass verschiedene Personen darauf mit Passwort und per TLS-Verschlüsselung zugreifen konnten. Für die Nutzung während der Studie wurde eine aktuelle Kopie der Daten auf die Instanz des Tablets gespielt, da dieses auch offline nutzbar sein musste.

Die Anzeigekomponente bestand aus dem Programm *Screen* und einer Website (GUI), die durch den *Screen* zur Verfügung gestellt wurde. Die Kommunikation zwischen *Screen* und Website war per WebSocket-Verbindung implementiert. Der *Screen* empfing die vom Empfehlungssystem ausgewählten Empfehlungen per XMPP-Nachricht und reichte sie für die Darstellung über die Website in eine Queue ein. Die benötigten Daten zur Darstellung wurden hierbei vom CMS abgerufen und auf der Website für die Darstellung zusammengefügt. Interaktionen, wie die Nutzerbewertungen oder die Abfrage der Stimmung des Nutzers, wurden von der Website per WebSocket an das *Screen*-Programm geschickt und anschließend per XMPP an die *Recommender*-Software weitergereicht. Protokoll-Daten wurden per XMPP an die *LoggerApp* gesendet.

Betriebssystem Alle Programme wurden auf dem Tablet unter Windows 8.1 genutzt. Da Projektpartner die Software möglichst problemlos mit ihren Systemen nutzen können sollten, wurden alle Programme so vorbereitet, dass sie bis auf wenige Ausnahmen, ohne Installation aus einem Ordner heraus, direkt ausführbar waren. Dieser Ordner konnte gepackt auf Systeme verteilt werden.

Die Telemetrie von Windows 8.1 wurde mittels vorhandener Datenschutzeinstellungen soweit wie möglich abgeschaltet. Automatische Updates, Virenschans sowie diverse Energiesparmaßnahmen (z. B. Standby, Abschalten des Bildschirms) wurden deaktiviert.

Das Startskript startete den verwendeten Browser (Google Chrome) im Vollbild (immer im Vordergrund) und der Windows Explorer wurde beendet (der Desktop war somit nicht mehr verwendbar). Hierdurch sollte ein versehentliches Beenden oder Verlassen der vom Browser dargestellten GUI des CARE-Systems verhindert werden.

¹³<https://www.drupal.de/> (10.12.2019)

¹⁴<https://www.apachefriends.org/de/index.html> (10.12.2019)

Empfehlungen

Der erste Prototyp verfügte über 48 Empfehlungen, wobei die meisten zur Verbesserung des physischen Wohlbefindens zählten. Die Empfehlungen wurden für die Studienteilnehmer weitestgehend personalisiert, um deren Relevanz zu erhöhen. Einige der Rätselaufgaben bezogen sich daher auf Augsburg, dem Wohnort der Teilnehmer. Zusätzlich wurden Fingerübungen für Senioren aus mehreren Quellen im Internet ausgewählt und sinngemäß im System umgesetzt, da die Studienteilnehmerin hieran besonderes Interesse zeigte.

9.4.2 Evaluation

Das System wurde mittels einer Feldstudie in einem Haushalt bestehend aus einem Seniorenpaar evaluiert. Die Frau war 76 Jahre, der Mann 75 Jahre alt. Die Studie lief 14 Tage lang. Zunächst musste eine geeignete Position für die Aufstellung des Systems in der Wohnung gefunden werden. Hierbei wurde zusammen mit den Bewohnern die Küche ausgewählt, da diese regelmäßig genutzt wurde und ausreichend Platz vorhanden war. Die Höhe des Displays wurde an die Körpergröße der Senioren angepasst. Die Studienteilnehmer wurden instruiert, das System nicht abzuschalten oder das Netzteil des Tablets zu entfernen. Die Nutzung des CARE-Systems wurde nicht weiter erklärt, da es so konzipiert war, dass es weitestgehend selbsterklärend nutzbar sein sollte. Während der Studie wurde die Funktionalität des Systems zweimal durch die Entwickler geprüft und die Daten bei den Besuchen gesichert. Das System lief ansonsten durchgängig. Den Senioren wurden zusätzlich zwei AX3-Armbänder gegeben und sie gebeten, diese so oft wie möglich zu tragen. Diese Armbänder wurden allerdings ausschließlich genutzt, um zu prüfen, ob sie bei zukünftigen Prototypen eingesetzt werden können. Hierzu war die interne Datenaufzeichnung der Bänder aktiviert. Abbildung 9.16 links zeigt die Studienteilnehmerin während der Interaktion mit dem CARE-Prototyp. Am rechten Handgelenk trägt sie das AX3-Armband.

Sensordaten

Es wurde zusätzlich eine Analyse der aufgezeichneten System-Protokoll-Daten durchgeführt. Diese Daten enthielten, den Zeitpunkt der Anzeige von Empfehlungen, eventuelle Bewertungen der Empfehlungen vom Nutzer, die nutzereigene Bewertung der Stimmung (siehe Abbildung 9.15), Sensordaten sowie automatisch aufgenommene Fotos vom System bei Erkennung einer Nutzerpräsenz. In Abbildung 9.16 rechts ist der Studienteilnehmer mit dem System dargestellt (siehe a)). Zusätzlich sind zwei der von ihm aufgezeichneten Fotos (siehe b) und d) sowie die zu diesem Zeitpunkt dargestellte Empfehlung abgebildet (siehe c)).



ABBILDUNG 9.15: Abfrage der Nutzerstimmung beim zweiten Prototyp (lediglich das Design unterscheidet sich gegenüber dem ersten Prototyp).

Mittels der aufgezeichneten Daten war es möglich festzustellen, ob Personen im Raum (Raumpräsenz) waren bzw. vor dem System standen (Nutzerpräsenz). Um herauszufinden, ob sich eine Person vor dem System befand, wurden die Daten des IR-Abstandssensors ausgewertet. Falls

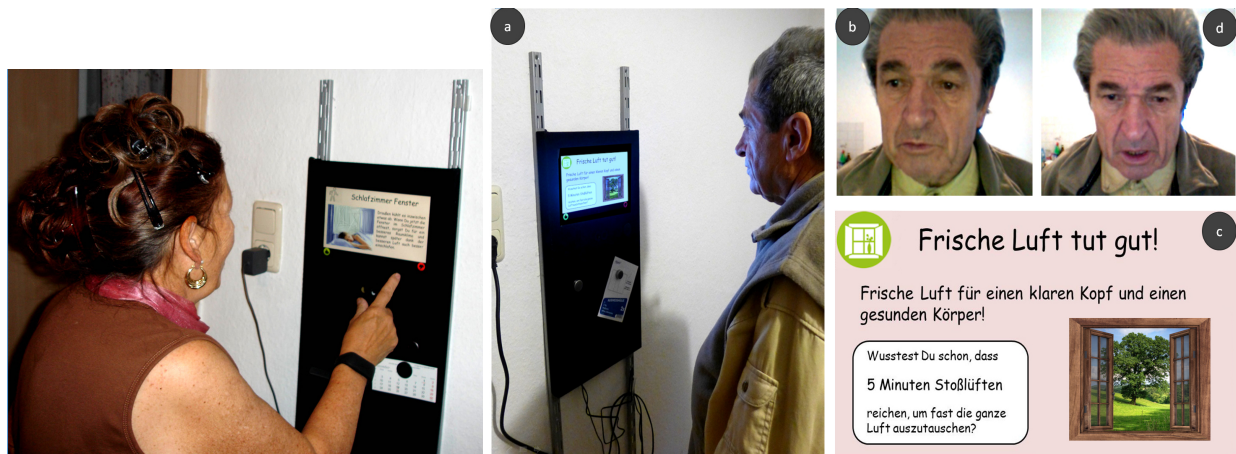


ABBILDUNG 9.16: Links: Aufgebauter Prototyp 1 mit Studienteilnehmerin. Die Probandin trägt einen AX3-Sensor am rechten Handgelenk (Abbildung angepasst aus (Seiderer, Hammer, André, Mayr & Rist, 2015)). Rechts: Prototyp 1 mit Studienteilnehmer, aufgezeichneten Fotos und der angezeigten Empfehlung (Abbildung angepasst aus (Rist, Seiderer, Hammer, Mayr & André, 2015)).

für mindestens zwei Sekunden ein Nutzer erkannt wurde (hohe Signalamplitude), wurde davon ausgegangen, dass die Person vor dem Bildschirm steht und den Inhalt wahrnimmt. Kurze Signalausschläge wurden als „Vorbeigehen“ eingestuft, wobei höchstwahrscheinlich der angezeigte Inhalt nicht wahrgenommen wurde. Während der Studie konnten 29 % der erfassten Präsenzen als Nutzerpräsenz vor dem System eingestuft werden. In Abbildung 9.17 oben ist die tägliche Verteilung der Nutzerpräsenz dargestellt. Es zeigt sich, dass diese vor allem zwischen 8 und 23 Uhr stattfanden. Die Daten des IR-Abstandssensors (Vorbeigehen) in Verbindung mit Informationen zur Raumhelligkeit, Audiointensität und Luftqualität wurden bei diesem Prototyp genutzt, um festzustellen, wann sich Personen im Raum befanden.

Die Studie wurde im November durchgeführt, sodass die Küche zwischen 8 und 16 Uhr durch einfallendes Tageslicht von zwei Fenstern beleuchtet war (siehe Abbildung 9.17 zweites Diagramm von oben). Außerhalb dieser Zeit muss künstliche Beleuchtung durch eine Person eingeschaltet worden sein. Zur Auswertung wurden aus den Lichtsensordaten Ereignisse für das Ein- (steigende Flanke) und Ausschalten (fallende Flanke) erzeugt. In den meisten Fällen wird das Einschalten des Lichtes bedeutet haben, dass eine Person den Raum betreten hat. Eine gewisse Ungenauigkeit war allerdings vorhanden, da das Licht beim Verlassen nicht immer ausgeschaltet wurde. In der Abbildung ist die tägliche mittlere Verteilung der Einschalt-Events dargestellt.

Um die Präsenz von Personen im Raum zu erkennen, waren auch die aufgezeichneten Audiointensitätsdaten (RMS) nutzbar. Da es sich um die Küche handelte, können diese von menschlicher Aktivität stammen, wie beispielsweise dem Abwasch. Allerdings konnten im Signal auch periodisch auftretende höhere konstante Intensitäten festgestellt werden, die dem Kühlschrank zuzuordnen waren. Daher wurde die Varianz und ein Schwellwert verwendet, um Nutzeraktivität abschätzen zu können. Ungenauigkeiten konnten allerdings durch Geräusche auftreten, die von außen, dem benachbarten Wohnzimmer oder dem ebenfalls in der Wohnung lebenden Hund stammten. Die mittlere Verteilung der ermittelten Audioaktivität ist in der dritten Zeile von oben in Abbildung 9.17 dargestellt.

Zusätzlich war beim CARE-Prototyp ein VOCs-Gassensor vorhanden. Dieser konnte auf von Lebewesen oder bei Aktivitäten wie z. B. Kochen verursachte Gerüche bzw. Gase reagieren. Immer wenn ein Anstieg der VOCs-Konzentrationen und somit eine Verschlechterung der Luftqualität festgestellt werden konnte, war eine Nutzerpräsenz im Raum wahrscheinlich, insbesondere falls kein Fenster geöffnet war. In der untersten Zeile von Abbildung 9.17 ist die mittlere Verteilung der Nutzerpräsenz im Raum basierend auf den Gassensordaten dargestellt.

Alle Nutzerpräsenzen zeigten ab 16 Uhr erhöhte Werte. Dies stimmte mit der Aussage der

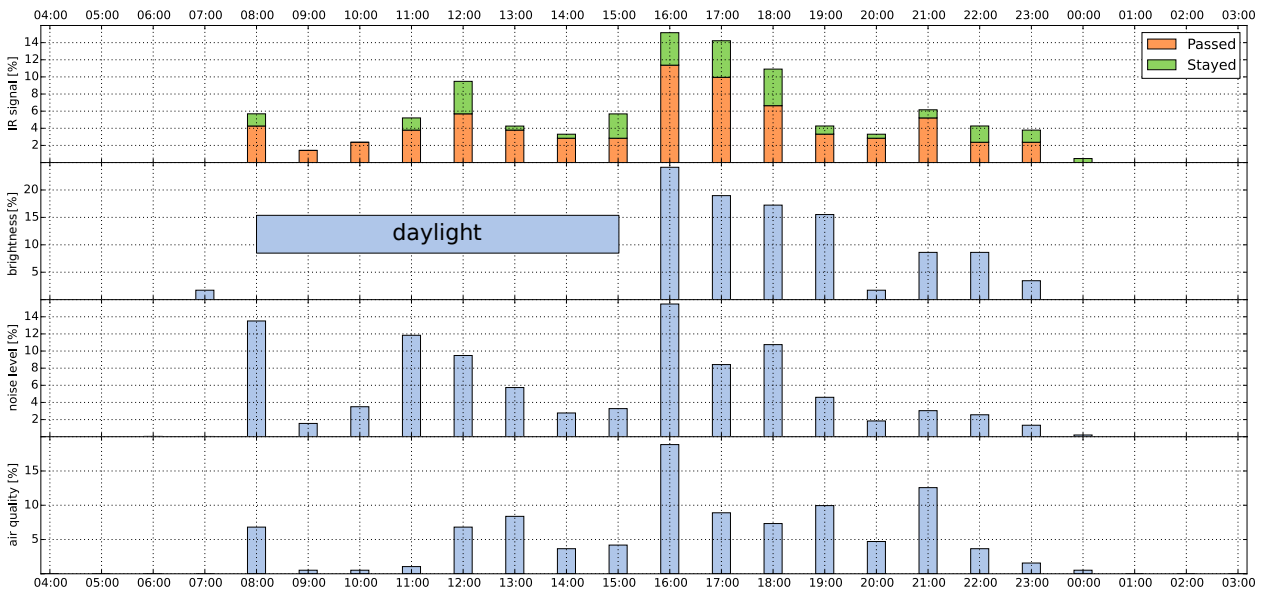


ABBILDUNG 9.17: Mittlere tägliche Verteilung der Nutzerpräsenz bestimmt aus verschiedenen Sensorsignalen des ersten CARE-Prototyps.

Senioren überein, dass sie zu dieser Zeit meistens in der Küche Speisen zubereiteten oder sie verzehrten. Generell bestanden zwischen den Nutzerpräsenzen, die von den IR-Abstandsdaten abgeleitet wurden und den anderen drei Datenquellen hohe Korrelationen, die bereits in der Abbildung visuell ersichtlich sind. IR-Abstand und Helligkeit: ($r=0,94$, $p<0,0005$), IR-Abstand und Audiointensität ($r=0,83$, $p<0,0005$) und IR-Abstand und Luftqualität ($r=0,82$, $p<0,0005$).

Wie eine Interaktion mit dem System entsprechend der Auswertung der Daten des IR-Abstandssensors, Lichtsensors sowie der Audiointensität nachvollziehbar ist, zeigt Abbildung 9.18 und wird im Folgenden exemplarisch erklärt. Um 22:59:27 Uhr betrat eine Person die Küche und schaltete das Licht an. Dies ist in der Abbildung bei „brightness“ ersichtlich. Da es sich um eine Energiesparlampe handelte, nahm die Helligkeit langsam zu. Ein kurzer Ausschlag beim IR-Abstandssensor-Signal (IR analog value) um 22:59:30 indiziert, dass die Person relativ nahe am CARE-System vorbeilief, eventuell um kurz in die Küche zu gehen. Ausschläge bei der Varianz der Audiointensität (noise level RMS var.) sind im untersten Diagramm dargestellt.

Bei der nächsten Erhöhung des IR-Abstandssensorsignals um 23:00:11 blieb die Person für längere Zeit vor dem System stehen. In diesem Fall wurde vom Bilderrahmen-Modus in den Empfehlungs-Modus umgeschaltet. Dies ist in der Abbildung ganz oben (Display) mit „P“ für „picture“ und „R“ für „recommendation“ gekennzeichnet. Da es bereits später am Abend war und sich die Nutzer höchstwahrscheinlich zum Schlafen vorbereiteten, wurden Empfehlungen herausgefiltert, die mit physischem, mentalem und sozialem Wohlbefinden zu tun hatten. Entsprechend wurden ausschließlich Empfehlungen aus den Kategorien umgebungsbezogenes und emotionales Wohlbefinden berücksichtigt. In diesem Fall war es nicht notwendig gewesen, das Raumklima beispielsweise durch Lüften zu verbessern, daher wurde eine Entspannungsübung ausgewählt.

Bei diesem Prototyp wurde eine bestimmte Zeitdauer genutzt, nach der in den Bilderrahmen-Modus zurückgeschaltet wurde, falls keine Bewertung zuvor abgegeben wurde. Die Empfehlung wurde in diesem Fall daher genau eine Minute angezeigt. In den IR-Abstandssensordaten ist ersichtlich, dass die Person noch ca. 1,5 Minuten vor dem System verweilte. Es ist nicht erkennbar, ob sie die Übung weiter durchführte oder das Bild betrachtete. Da dies jedoch öfters auftrat, ist es wahrscheinlicher, dass die Übungen weiter durchgeführt wurden.

Um 23:02:48 verließ die Person den Erkennungsbereich des IR-Abstandssensors für 12 Sekunden. Da eine Audioaktivität feststellbar war, war die Person noch in direkter Nähe. Um 23:03:03 erfasste der IR-Abstandssensor die Person erneut. Es wurde wiederum eine Empfehlung ausgewählt, wobei

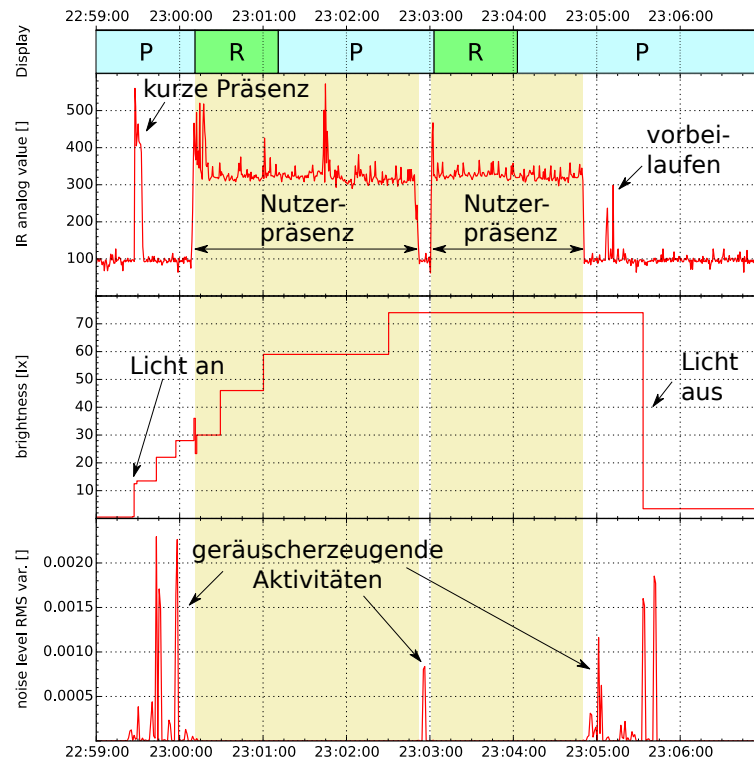


ABBILDUNG 9.18: Ausschnitt aus den aufgezeichneten Sensordaten während einer Nutzerinteraktion mit dem ersten CARE-Prototyp.

es sich diesmal um eine andere Entspannungsübung handelte. Die Person verweilte diesmal für ca. zwei Minuten vor dem Prototyp. Wiederum wurde zuvor vom System in den Bilderrahmen-Modus zurückgeschaltet und die Übung von der Person augenscheinlich weiter durchgeführt.

Die Person verließ um 23:04:51 endgültig den Erfassungsbereich des IR-Abstandssensors und führte anschließend offenbar eine Aktion im Raum aus, die Geräusche produzierte. Um 23:05:40 wurde das Licht ausgeschaltet und eine Audioaktivität war nicht mehr feststellbar. Ab diesem Zeitpunkt wurde der Raum offensichtlich verlassen.

Nutzerbefragung

Neben der Datenanalyse wurde eine Befragung mittels eines strukturierten Interviews durchgeführt. Hierzu wurden die Heuristiken von Mankoff et al. (2003) zur Evaluation von ambienten Anzeigen und die von Kientz et al. (2010) zur Evaluation persuasiver Gesundheitstechnologien eingesetzt. Die übersetzten Namen der Heuristiken sind teilweise als Titel der folgenden Abschnitte angegeben.

Nützliche und relevante Information Die Studienteilnehmer wurden gefragt, an welche Empfehlungen sie sich erinnern können. Sie erwähnten hierbei die physischen und Meditationsübungen sowie Knobelaufgaben. Das Paar konnte sich nicht an Empfehlungen bezüglich der Verbesserung des Raumklimas erinnern. Möglicherweise war es den Senioren unangenehm, darüber zu sprechen.

Die beiden Teilnehmer besaßen unterschiedliche Präferenzen bezüglich der Empfehlungskategorien. Die Seniorin bevorzugte die physischen Aufgaben für die Finger und Hand. Zusätzlich empfand sie die Entspannungsübungen als nützlich. Diese Art von Übungen waren ihr bereits bekannt, allerdings hatte sie diese bisher nicht selbst ausprobiert. Sie erwähnte hierbei, dass sie gerne längere Übungen (ca. 20 bis 30 Minuten) zweimal am Tag, regelmäßig durchführen wollen

würde. Dahingegen präferierte der Senior die Witze. Beide Teilnehmer betonten, dass eine größere Anzahl an Inhalten vorhanden sein müsste.

Geeignete Zeit und geeigneter Ort Beide Studienteilnehmer stimmten darin überein, dass die Küche ein geeigneter Ort für das CARE-System war, da sie hier häufig am System vorbeikamen. Die Seniorin wünschte sich allerdings, morgens und abends längere Interaktion mit dem System, um mehrere Übungen hintereinander durchführen zu können. Der Senior erwähnte hierbei keine Probleme, was daran liegen konnte, dass er weniger an den Übungen interessiert war.

Einfachheit der Nutzung Beide Senioren waren derselben Meinung, dass die Bilder auf dem Tablet gut erkennbar und die Texte gut lesbar waren. Sie empfanden die Informationsmenge der Anzeigen passend. Die Bedeutung der Buttons war ihnen klar. Obwohl eine direkte Interaktion mit dem System nur eingeschränkt stattfand, konnten einige Usability-Probleme festgestellt werden. Der IR-Abstandssensor erfasste beispielsweise nicht immer zuverlässig Personen vor dem System oder die Verarbeitung benötigte zu viel Zeit. Da der Erkennungsbereich des Sensors für die Nutzer nicht ersichtlich war, konnte es zudem vorkommen, dass sich ein Nutzer unbeabsichtigt aus dem Erfassungsbereich bewegte und somit vorzeitig in den Bilderrahmen-Modus zurückgeschaltet wurde.

Flexibilität und Effizienz der Nutzung Um die Studienteilnehmer nicht zu überfordern, wurde das Nutzerinterface bewusst einfach gestaltet. Es wurde daher beispielsweise keine Navigationsmöglichkeit integriert. Dennoch wünschten sich beide Senioren mehr Möglichkeiten bei der Steuerung. Die Nutzerin erwähnte zudem, dass sie für gewisse Übungen Gegenstände benötigte, z. B. einen Igelball, die sie erst holen musste, wodurch das System ungewollt in den Bilderrahmen-Modus zurückschaltete. Dies geschah teilweise auch, wenn sie spontan etwas anderes machen musste. Ein gezieltes Ansteuern einer zuvor angezeigten Empfehlung war jedoch nicht möglich. Sie wünschte sich daher, insbesondere bei physischen Aufgaben, eine Möglichkeit das System zu pausieren.

Peripherie der Anzeige Obwohl das Display dauerhaft eingeschaltet war, wurde es von den Senioren nicht als störend empfunden. Das Display wurde somit peripher wahrgenommen. Dies steht im Gegensatz zu Studien von Consolvo und Towle (2005), deren ambientes Display in dunklen Lichtsituationen als störend empfunden wurde. In der eigenen Studie konnte dieses Problem durch die Platzierung des Systems in der Küche verhindert werden. Eine Positionierung in der Nähe eines Fernsehgerätes oder im Schlafzimmer wäre daher nicht geeignet.

Nutzerausbildung Die Nutzer sollten verstehen, warum die Aktionen für ihr Wohlbefinden vorteilhaft sind. Die Seniorin verfügte über ein klares Verständnis hinsichtlich des Zwecks der Übungen. Ihr Wissen stammte von der Teilnahme an einem Rehabilitationsprogramm und ihrem früheren Beruf als Altenpflegerin. Die kurzen Erklärungen bei den Empfehlungen waren daher für die Seniorin ausreichend. Eine überzeugende Gestaltung von Empfehlungen wurde allerdings im Rahmen des CARE-Projekts nicht untersucht. In einer späteren, weiterführenden Arbeit von Hammer et al. (2016) wurde die Überzeugungskraft von Formulierungen für Empfehlungen sowie die Nutzung von Robotern in Verbindung mit Senioren erforscht.

Angemessene Funktionalität Beide Senioren merkten an, dass sich das System gut in ihre tägliche Routine einfügen ließ. Insbesondere die Seniorin war vom CARE-System beeindruckt. Sie sagte, dass CARE ihr dabei half, aktiv zu bleiben und nicht träge zu werden. Sie schien allerdings generell eine höhere intrinsische Motivation als ihr Partner zu besitzen. Dies könnte daran gelegen haben,

dass die angezeigten Inhalte eher ihren persönlichen Bedürfnissen entsprachen. Zusätzlich besaß die Seniorin ein besseres Verständnis hinsichtlich des Zwecks der Empfehlungen.

Zusammenfassung Nach der Studie zeigte das Feedback der Senioren, dass sich das CARE-System gut in den Alltag integrieren ließ. Insbesondere kamen bei der Seniorin die physischen Übungen gut an. Allerdings konnten Probleme beim zeitlichen Ablauf der Empfehlungsanzeigen beobachtet werden. Beispielsweise hätte die Seniorin gerne regelmäßig längere Übungen im Zeitrahmen von 20 bis 30 Minuten durchgeführt, was bei der Implementierung des ersten Prototyps nicht berücksichtigt worden war. Die Senioren wünschten sich zudem mehr Kontrolle über das Nutzerinterface.

Abgeleitete Anforderungen für den zweiten Prototyp

Aus der Datenanalyse sowie den Nutzerbefragungen konnten verschiedene Anforderungen für weitere Prototypen abgeleitet werden. Nachfolgend werden diese zusammengefasst.

Den Nutzern mehr Kontrolle geben Die Interaktionsmöglichkeiten waren beim ersten Prototyp bewusst einfach gehalten und sollten bevorzugt mittels impliziter Interaktion möglich sein. Es zeigte sich allerdings, dass die Studienteilnehmer mehr Kontrolle über den zeitlichen Ablauf der Empfehlungen wünschten.

Verbesserung der Usability des berührungssensitiven Interfaces Um die Empfehlungen in späteren Versionen des Systems stärker personalisieren zu können, wurden bereits Bewertungen vom Nutzer abgefragt. Zudem wurde die Nutzerstimmung über das GUI erfasst. Obwohl die Studienteilnehmer aussagten, dass sie die Bedeutung der Buttons verstanden hätten, wurde nur selten Feedback gegeben. Durch die Auswertung der aufgezeichneten Accelerometer-Daten des Tablets konnte allerdings festgestellt werden, dass nicht alle Berührungen vom Touchscreen registriert wurden, obwohl leichte Erschütterungen in den Beschleunigungsdaten erkennbar waren. Möglicherweise war der Touchscreen für ältere Menschen mit relativ trockenen Fingerkuppen nicht ausreichend empfindlich, sodass entsprechend zur Arbeit von Leonardi et al. (2010) vergleichbare Eingabeprobleme auftraten. Der Einsatz eines Tablets mit empfindlicherem Touchscreen sowie die Nutzung von deutlicherem Systemfeedback beim Berühren der angezeigten Buttons, können dieses Problem reduzieren.

Integration von Anreizen für Nutzer Beim ersten Prototyp waren nur wenige Anreize enthalten, die Nutzer motiviert hätten, Empfehlungen zu beherzigen. Beispielsweise ist es von Vorteil, wenn Nutzer beobachten können, wie weit ihr persönlicher Fortschritt bei gesetzten Zielen ist. Hierzu können Konzepte aus der Gamifizierung sowie sozialen Medien einbezogen werden.

Verbesserung der Kontextsensitivität von CARE Um die Empfehlungen zum Kontext passender auswählen zu können, war es notwendig, die Sensorik weiter auszubauen und die Bewertungen der Empfehlungen zu berücksichtigen.

Erhöhung der Empfehlungsanzahl Um zu erreichen, dass Nutzer das CARE-System über einen längeren Zeitraum nutzen, sollte die Anzahl der Empfehlungen erhöht werden. Zusätzlich kann das soziale Wohlbefinden stärker berücksichtigt werden, indem beispielsweise Familienmitglieder an das System Nachrichten schicken können. Hierdurch ist es gemäß Cornejo et al. (2013) möglich, die Eingebundenheit von Senioren zu erhöhen, und kann daher zudem die Attraktivität des Systems verbessern.

9.5 Entwicklung CARE-Prototyp 1.5 und 2

Aus den Erkenntnissen der Entwicklung und Evaluation des ersten Prototyps wurden zwei weitere Prototypen entwickelt. Der Prototyp 1.5, die CARE-Box, war hierbei ein Zwischenschritt zu Prototyp 2, wobei die CARE-Box als portabler Demonstrator angedacht war. Der Prototyp 2, der, wie der erste Prototyp in einer Feldstudie evaluiert wurde, war wiederum für eine Integration in die Wohnumgebung vorgesehen und umfasste mehr Sensorik als die CARE-Box. Auf beiden Prototypen konnte allerdings dieselbe Software ausgeführt werden, da identische Tablets genutzt wurden. Teile dieses Abschnitts wurden in der eigenen Arbeit (Rist, [Seiderer](#) & André, 2018) veröffentlicht.

9.5.1 Weiterentwicklungen

Gemäß den aus der Evaluation des ersten Prototyps abgeleiteten Anforderungen wurden umfangreiche Anpassungen am CARE-System durchgeführt.

- Die **Sensorik und Kontexterfassung** wurden deutlich ergänzt, überarbeitet und verbessert. Latenzen wurden durch Optimierungen und durch Nutzung leistungsfähigerer Tablet-Hardware deutlich reduziert.
- Beim **Empfehlungssystem** wurden die Bewertungen und weitere Kontextinformationen berücksichtigt sowie immer mehrere Empfehlungen pro Nutzerinteraktion ausgewählt.
- Das **Nutzerinterface** wurde verbessert und erweitert. Es war nun zusätzlich möglich, zwischen einzelnen Empfehlungen zu navigieren (zurück / vor), zu pausieren und zu bestätigen, ob eine Empfehlung durchgeführt wurde (falls keine automatische Erkennung vorhanden war). Ein Menü erlaubte es, einen Hilfetext anzuzeigen sowie gezielt auf bestimmte Empfehlungskategorien zuzugreifen, um beispielsweise alle Übungen ansehen und über längere Zeiträume verschiedene Übungen durchführen zu können. Das Feedback der GUI-Buttons wurde verbessert, indem ein kurzer Ton bei erkanntem Tastendruck abgespielt wurde und zusätzlich ein eingedrückter Knopf, während des Drückens, dargestellt wurde. Des Weiteren wurden Hardware-Buttons mit LED-Beleuchtung hinzugefügt, die optional nutzbar waren, falls Probleme mit der Eingabe per Touchscreen auftraten.
- Außerdem wurde ein **Postkartenmodus** ergänzt, bei dem von Verwandten der Nutzer per Android-App oder Webinterface gesendete kurze Texte inklusive eines Bildes vom System dargestellt werden konnten. Bei Nutzerpräsenz wurden neue Postkarten vor den Empfehlungen angezeigt. Über das Menü konnte durch alle empfangenen Postkarten navigiert werden.
- Als weitere Systemausgabe wurde **ambientes Licht** integriert, das eine Vorschau darauf geben sollte, welche Wohlbefindenskategorie vom System ausgewählt oder ob neue Postkarten empfangen wurden. Es wurde hierfür eine RGB-LED in der Nähe des CARE-Systems platziert, die beim Betreten des Raumes für eine bestimmte Zeitdauer aktiviert wurde.

Generell konnte die Robustheit des Systems verbessert werden. Zusammen mit den anderen Anpassungen wurden somit die wesentlichen Probleme des ersten Prototyps berücksichtigt und weitestgehend beseitigt.

9.5.2 Umsetzung

Hardware

Als wesentliche Neuerung wurde beim CARE-Prototyp 1.5 und 2 das HP Omni-10-Tablet durch ein leistungsfähigeres Microsoft Surface-3-Pro-Tablet ersetzt (CPU: Intel Core i5-4300U, RAM: 4 GiB, Speicher: 128 GiB) mit einem größeren und berührungsempfindlicheren Touchscreen (12 Zoll 3:2

Format bei einer Auflösung von 2160×1440 Pixeln). Es kostete ca. 1000 €. Dies reduzierte bereits die Latenzen des Systems und dank des hier nutzbaren 64-Bit-Systems war die Datenbankgröße nicht mehr auf 2 GiB begrenzt. Zusätzlich stellte sich bei Tests heraus, dass der Touchscreen bereits bei leichter Berührung reagierte, was beim vorherigen HP Tablet für Senioren ein Problem darstellte.

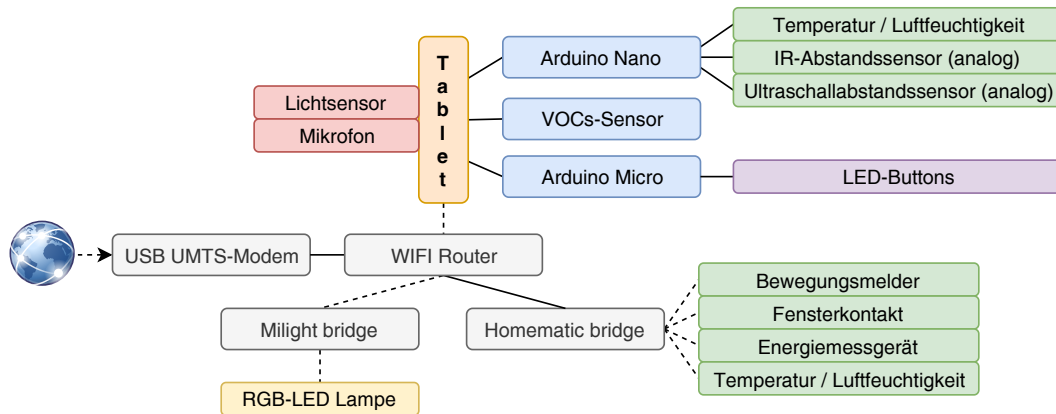


ABBILDUNG 9.19: Hardware-Aufbau CARE-Prototyp 1.5 / 2. Gestrichelte Linien waren Funkverbindungen.

In Abbildung 9.19 ist der Hardware-Aufbau des CARE-Prototyps 1.5 / 2 dargestellt. Die Verwendung der integrierten Tablet-Sensoren beschränkte sich bei diesem Prototyp auf Lichtsensor und Mikrofon. Ansonsten waren per USB ein Arduino Nano, Arduino Micro und der AMS iAM VOCs-Gassensor angeschlossen. Der Arduino Nano diente dem Auslesen eines Temperatur- / Luftfeuchtigkeitssensors (DHT22), IR-Abstandssensors (Sharp GPY0A02YK0F) und Ultraschall-Abstandssensors (HC-SR04) mit einer Abtastrate von 4 Hz. Der Arduino Micro erfasste den Tastendruck von LED-Tastern (mit Software-Entprellung) und konnte zudem per serieller Schnittstelle die jeweiligen LEDs der Taster an- bzw. ausschalten. Der Arduino Micro verfügte über einen ATmega32U4, der direkt an der USB-Schnittstelle angeschlossen war und eine Tastatur simulierte, wobei jeder Taster einen anderen Buchstaben auslöste. Dies erlaubt eine direkte Erfassung der Eingaben durch beliebige Anwendungen. Eine Tastenkombination aus drei Tasten konnte den Kiosk-Modus des Webbrowsers beenden.

Das Tablet war per WLAN mit einem Router verbunden, der die Möglichkeit bot, per USB-Stick ein UMTS-Modem anzusteuern, um eine Internetverbindung in Wohnungen zu ermöglichen, in denen ansonsten keine zur Verfügung gestanden hätte. Der UMTS-Stick wurde mit einem USB-Verlängerungskabel angeschlossen, um eine empfangsstarke Position wählen zu können. Mit dem WLAN war zudem die Milight-Bridge verbunden, die zur Ansteuerung einer RGB-LED-Lampe benötigt wurde. Des Weiteren war per LAN-Kabel eine Homematic-Bridge an den Router angeschlossen, um die Sensorik des Homematic-Systems nutzen zu können. Für das System waren ein Bewegungsmelder, Fensterkontakt, Energiemessgerät und Temperatur- / Luftfeuchtigkeitssensor vorgesehen. Der Bewegungsmelder wurde zur Überwachung des Raumes genutzt. Der Fensterkontakt konnte verwendet werden, um festzustellen, ob gelüftet wurde. Das Energiemessgerät war ein Zwischenstecker, der zwischen die Steckdose und den Netzstecker des Fernsehers gesteckt werden sollte, um indirekt erkennen zu können wie lange ferngesehen wurde und somit Informationen über die Nutzeraktivität zu erhalten.

Der einzige Unterschied zwischen der CARE-Box (Prototyp 1.5) und Prototyp 2 bestand darin, wie das Tablet, die Sensorik und zusätzliche Hardware verbaut wurden. Bei der CARE-Box (siehe Abbildung 9.20 links) handelte es sich um eine weiß gestrichene Holzbox mit einem Deckel, der mit zwei Scharnieren verbunden war. Ziel war es, ein CARE-System mit allen wesentlichen Sensoren transportabel zu realisieren. Es war primär für Demonstrationen gedacht. Das Tablet wurde in eine Aussparung im Deckel eingehängt, wobei weiche Filzplättchen eine Beschädigung verhinderten. Der Ultraschall-Abstandssensor und die Taster waren in der Front verbaut. Der IR-Abstandssensor



ABBILDUNG 9.20: Die transportable CARE-Box (Prototyp 1.5). Links ist die Außenansicht mit optionaler RGB-Lampe für ambientes Licht dargestellt. Das Tablet wird von oben eingelegt, verkabelt (USB + Strom) und mit Filzplättchen fixiert. Rechts die Elektronik innerhalb der Box zur Ansteuerung der LED-bleuchteten Taster und Erfassen von Sensordaten. Innerhalb der Box sind ebenfalls eine Steckdosenleiste, das Netzteil des Tablet-PCs sowie ein USB-Hub enthalten.

war im Deckel enthalten. Die restliche Sensorik und Hardware konnten innerhalb der Box platziert werden (siehe Abbildung 9.20 rechts), in der sich auch ein Steckdosenverteiler für die Netzteile befand. Die Box benötigte daher nur eine externe 230-V-Steckdose, um betrieben zu werden. Falls die ambiente Beleuchtung eingesetzt werden sollte, konnte die LED-RGB-Birne in eine vorhandene Lampe mit E27-Fassung geschraubt oder der weiße, halbtransparente Lampenschirm mit E27-Fassung (siehe Abbildung 9.20 links) verwendet werden.

Der Prototyp 2 wurde in ein Regal eingebaut. Abbildung 9.21 links zeigt den Aufbau während der Evaluation in der Wohnung der Studienteilnehmerin, rechts ist eine spätere Integration in ein anderes Regal abgebildet. Auf dem linken Regal befinden sich oben die in der Studie verwendete Leuchte, der Homematic-Bewegungsmelder (links neben der Lampe) und die Homematic-Bridge sowie der VOCs-Gassensor (rechts neben der Lampe) dargestellt. Der Router und wesentliche sonstige Teile der Elektronik waren hinter dem Tablet platziert. Unterhalb der beleuchteten Taster waren der Ultraschall- und IR-Abstandssensor nebeneinander angebracht. Im Regal befand sich ein blauer Igelball der Seniorin für bestimmte physische Übungen. Auf dem Foto sind im Regal der zum Aufnahmezeitpunkt noch nicht installierte Homematic-Zwischenstecker für die Energiemessung und Homematic-Temperatur- / Luftfeuchtigkeitssensor ebenfalls sichtbar.

Abbildung 9.21 rechts zeigt den Demonstrator bei schwacher Umgebungsbeleuchtung. Hierbei sind die LEDs der Taster gut sichtbar und erleichterten deren Bedienung auch bei ungünstigen Lichtverhältnissen. Nur die aktuell aktivierten und auf dem Touchscreen dargestellten Buttons waren beleuchtet. Eine eindeutige Zuordnung der Taster und der Bildschirmdarstellung wurde durch die Position, Farbe sowie Beschriftung ermöglicht. Von links nach rechts betrachtet haben die Taster folgende Bedeutung, die über das Menü auch als Hilfe angezeigt werden konnte: Zurück, pausieren, vorwärts, Menü, „Aufgabe erledigt“, negative und positive Bewertung. Die Symbole der Buttons für die Navigation orientierten sich an Fernbedienungen für TV-Geräte mit Abspielfunktion von Audio oder Videodateien. Die Buttons zur Navigation waren weiß, für das Menü blau, für „Aufgabe erledigt“ orange und zur Bewertung rot bzw. grün gefärbt. Die Gruppierung bestimmter Taster, z. B. zur Navigation, wurde durch deren geringeren Abstand zueinander zusätzlich verdeutlicht (Gesetz der Nähe).

Software

Überblick Die verwendete Software für die Prototypen 1.5 und 2 war identisch. Im Vergleich zum ersten Prototyp fand eine generelle Reduktion der Anzahl an Einzelprogrammen statt, was zu einer



ABBILDUNG 9.21: Links: Prototyp 2 Aufbau in der Wohnung der Studienteilnehmerin. Mitte: Integration des Prototyps in ein Regal. Rechts: Prototyp 2 bei schwacher Umgebungsbeleuchtung. Aktivierte Buttons sind auf dem Bildschirm sowie durch die LED-Beleuchtung der Taster sichtbar. Durch identische Position, Farbe und Symbole können die Software- und Hardware-Buttons eindeutig zugeordnet werden.

zusätzlichen Beschleunigung des Systems und erleichtertem Auffinden von Fehlern führte. Die Grundstruktur sowie das Datenformat blieben identisch. In Abbildung 9.22 ist der finale Softwareaufbau ersichtlich. Die Dateninterpretation (*Interpreter*), das Empfehlungssystem (*Recommender*) sowie die Ausgabekomponente (*Screen*) wurden in einem parallelisierten Java-Programm vereint und mittels eines asynchronen Eventbus (RxJava), anstelle von XMPP, miteinander verbunden. Sensordaten wurden nun primär mittels des leichtgewichtigeren MQTT-Protokolls vom integrierten MQTT-Server empfangen. Die Nutzung von XMPP war optional weiterhin möglich. Die MongoDB- und MySQL-Datenbank wurden durch PostgreSQL ersetzt. Es wurde ein einzelnes Datenbanksystem genutzt, da hierdurch die Systemkomplexität verringert werden konnte und PostgreSQL bei diesem System keine relevanten Geschwindigkeitsunterschiede im Vergleich zu MongoDB zeigte, dafür allerdings auch Relationen abbilden konnte, die mit einzelnen SQL-Abfragen verwendbar waren. In der Datenbank wurden Tabellen für die Sensordaten („Sensordata“), die berechneten Kontextinformationen („Context“), die Daten für das Nutzerprofil („Cases“) und die Postkarten („Content“) verwendet. Im Nutzerprofil wurden z. B. die Bewertungen für die Empfehlungen gespeichert. Zusätzlich konnte die Struktur der MySQL-Datenbank des Drupal-CMS nachgebildet werden („Content“). Das CMS musste auf dem Tablet nicht mehr während der Laufzeit verwendet werden und es konnten beim ersten Start des Systems alle Inhalte von einem über das Internet erreichbaren Drupal-CMS für eine Sprache lokal gesichert werden. Da in das Java-Programm auch ein minimaler Webserver mit WebSocket-Unterstützung integriert wurde, der auch für die HTML5-GUI (*Screen GUI*) des CARE-Systems zuständig war, konnte auf einen lokalen Apache-Server für das CMS verzichtet werden.

Als neues externes Programm wurde eine SSI-Instanz eingeführt, die das Interfacing und die Vorverarbeitung der Sensorik des Arduino Nano durchführte. Hierfür wurden mehrere neue

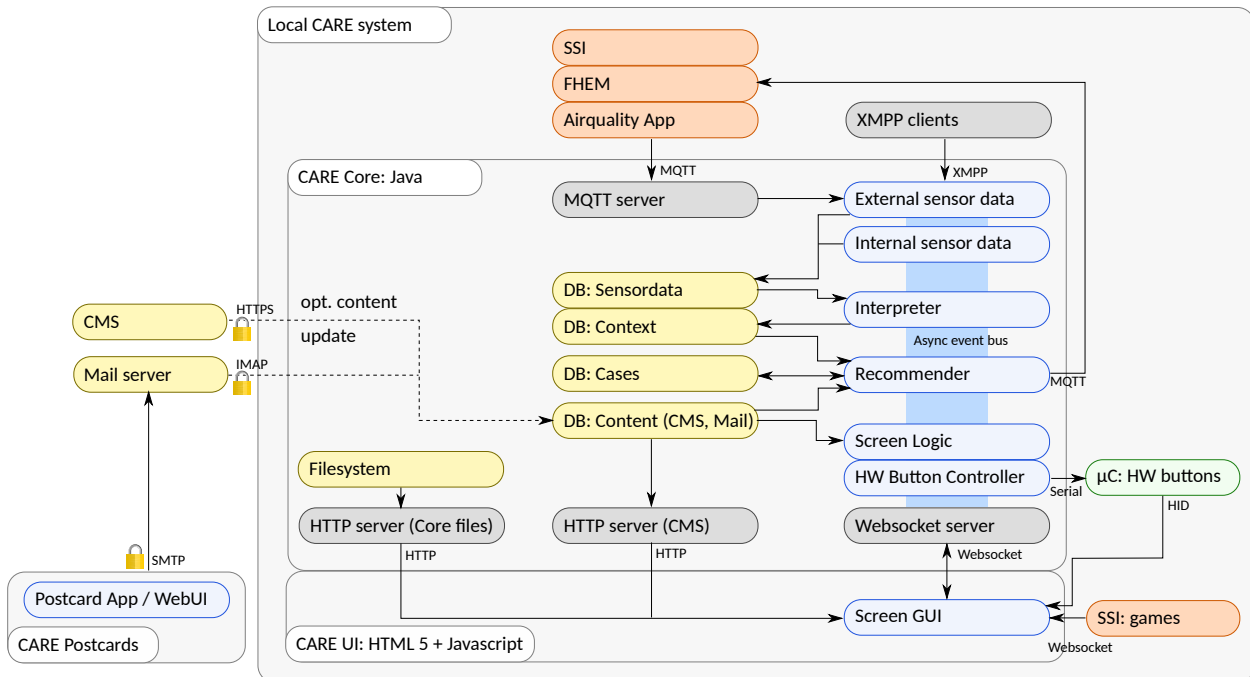


ABBILDUNG 9.22: Überblick über die Systemkomponenten des finalen CARE-Systems.

Plugins entwickelt (siehe Anhang C): *generic_serial* (Erfassung von Daten per serieller Schnittstelle), *winsensors* (Erfassung von Sensordaten über die Windows-Sensor-API), *mqtt* (Versenden von Daten mit dem MQTT-Protokoll), *eventinterpreter* (Verarbeitung von SSI-Events mit Lua-Skripten; in diesem Fall: Konvertierung in JSON-Format und Ergänzung von Informationen über Raum und Sensor). Ansonsten waren wie beim ersten Prototyp FHEM sowie die *Airquality App* im System enthalten, die nun allerdings per MQTT eingebunden wurden. FHEM steuerte nun die Hometric-Geräte sowie die Milight-Lampe an. Zur Ansteuerung der RGB-Lampe konnte das Empfehlungssystem (*Recommender*) die Befehle per MQTT direkt an FHEM senden.

Zur Interaktion mit bestimmten Übungen wurde bereits vorgesehen, dass SSI direkt Daten an die Logik der Übungen senden konnte (*SSI: games*). Hierfür wurde für SSI das *websocket*-Plugin (siehe Anhang C) implementiert. Die Logik einzelner Übungen konnte innerhalb von SVG-Vektorgrafiken enthalten sein, die vom CMS als Bild angesehen wurden, allerdings auch Javascript-Code enthalten konnten, der vom Webbrowser ausgeführt werden konnte. Diese Funktionalität wurde allerdings erst bei den späteren Untersuchungen zur Gamifizierung eingesetzt.

Eine Komponente im Java-Programm steuerte per serieller Schnittstelle die LEDs der Taster, passend zur Visualisierung, an.

Eine weitere Neuerung war die Möglichkeit, Postkarten an die Nutzer des Systems zu senden. Dies wurde mittels E-Mails implementiert, die neben dem Text als Anhang ein Bild enthalten konnten. Das Hauptprogramm des CARE-Systems überwachte diesen E-Mail-Account mittels des IMAP-Protokolls, wodurch es eventbasiert reagieren konnte, sobald eine neue Nachricht eintraf. Da auch Spam in das Postfach gelangen konnte, wurden Nachrichten nur mit einem bestimmten Betreff und festgelegter Struktur vom System weiterverarbeitet. Es war zudem nicht vorgesehen, dass Verwandte direkt E-Mails an das System versenden, da hierbei die Formatierung durch den Absender beachtet hätte werden müssen. Daher wurden ein Webinterface sowie eine Android-App entwickelt, die mit den Account-Daten die E-Mails an den Account selber versandten. Abbildung 9.23 links zeigt die GUI der Android-App. Auf der rechten Seite ist ein Beispiel für eine Postkarte abgebildet, wie sie vom CARE-System dargestellt worden ist. Es konnten der Absender, Empfänger, die eigene E-Mail-Adresse, eine Nachricht sowie ein Bild eingegeben werden. Der Absender, die Nachricht und optional ein Bild wurden vom System dargestellt. Der Empfänger

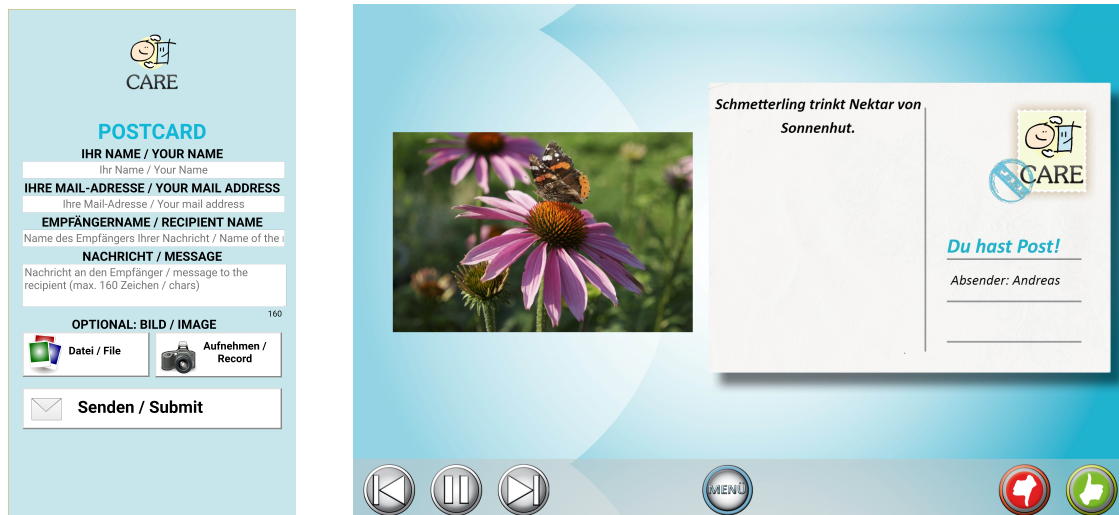


ABBILDUNG 9.23: Links: GUI der CARE Android-App zum Versenden von Postkarten an das CARE-System per E-Mail. Rechts: ein Beispiel für eine dargestellte Postkarte.

wäre primär bei einer Verwendung mit mehreren Nutzern von Relevanz gewesen und wurde daher für die geplante Studie mit einer Person von der Software ignoriert. Die eigene E-Mail-Adresse hätte grundsätzlich genutzt werden können, um Rückmeldungen vom System zu versenden. Diese Funktion wurde allerdings für diesen Prototyp nicht aktiviert, da die Postkarten primär als Gesprächsstoff für Telefongespräche dienen sollten.

Kontextinformationen und Empfehlungsauswahl Die Interpretation der Sensordaten wurde vollständig überarbeitet. Es waren dieselben Kontextinformationen des ersten Prototyps weiterhin enthalten, sie wurden allerdings optimiert und ergänzt. Eine Java-Komponente war nun zuständig, die Daten in der Datenbank einzufügen und mittels mehrerer Groovy-Skripte Kontextinformationen aus den Rohdaten zu gewinnen, die nun auch aufeinander aufbauen konnten. Dies war grundsätzlich vergleichbar mit einem Regelsystem bei dem Daten-Interpretationen „Fakten“ genannt werden. Ein wesentlicher Unterschied bestand allerdings darin, dass der Rete-Algorithmus von Forgy (1989) nicht eingesetzt wurde, den Regelsysteme meist nutzen. In diesem Fall wurden durch eine relativ geringe Anzahl an benötigten Interpretationen keine Konflikterkennung oder Schlussfolgerungen benötigt. Ein vollständiges Regelsystem zur Interpretation der Rohdaten hätte das System verlangsamen und zusätzlichen Arbeitsspeicher benötigen können.

Die möglichen Interpretationen zur Berechnung der Kontextinformation „Luftqualität“ (air quality) aus den Daten des VOCs-Sensor waren beispielsweise „schlecht“ (bad), „neutral“ (neutral) und „gut“ (good). Die Berechnung der Kontextinformationen wurde wie beim ersten Prototyp nur ausgelöst, sobald ein neuer Sensorwert vom System empfangen und in die Datenbank gespeichert wurde. Jede Kontextinformation (z. B. die Luftqualität) aus einer Dateninterpretation wurde zunächst im Speicher gehalten und erst in die Datenbank geschrieben, wenn alle evtl. voneinander abhängigen Interpretationen ausgeführt wurden. Andere Skripte konnten bereits berechnete Kontextinformationen direkt aus dem Speicher auslesen und lösten daher nur beim Fehlen einer Information deren Berechnung aus. Es konnte ein Zeitintervall festgelegt werden, nach dessen Ablauf Kontextinformationen neu interpretiert wurden. Zum Beispiel war es bei der sich nur langsam verändernden Raumtemperatur nicht notwendig, sie jede Sekunde neu zu interpretieren. Falls das Zeitintervall noch nicht überschritten war, wurde direkt das letzte Interpretationsergebnis aus dem Speicher zurückgegeben. Diese Optimierung konnte die Latenzen des Systems stark verringern.

```

boolean _calc(InterpreterEngine engine, GroovyInterpreter interpreter) {
    //get all sensors of type "airQuality"
    Vector <Sensor> airSensor = engine.getDbContext().getSensors(SensorTypes.airQuality);

    //did we find a sensor?
    if (airSensor != null && airSensor.size() > 0) {

        //do for each found sensor
        for (Sensor s : airSensor) {

            //get sensor data of the last 30 minutes
            DescriptiveStatistics stats = engine.getDbContext().getSensordataStatistics(s, 1800);

            //did we get sensor data?
            if (stats != null ) {

                //calculate the mean value
                double meanVal = stats.getMean();
                stats.clear();

                //rate the airquality
                String airRating = roomClimateAirQualityRating(meanVal);

                //get the room of the sensor
                Room room = engine.getDbContext().getRoom(s.get_id_room());

                //save the result as context element in memory "roomAirQuality"
                interpreter.addResult(airRating, (String[]) [interpreter.getName()].toArray(), (Sensor[])
[s].toArray(), (Room[]) [room].toArray(), null, null, ContextTypes.roomAirQuality, meanVal, false);
            }
        }
        airSensor.clear();
    }
    return true;
}

String roomClimateAirQualityRating( double voclevel) {
    if (voclevel < 1000) return ContextValues.RATINGS_THREE_STR.get("STR_GOOD");
    if (voclevel < 1500) return ContextValues.RATINGS_THREE_STR.get("STR_NEUTRAL");
    else return ContextValues.RATINGS_THREE_STR.get("STR_BAD");
}

```

ABBILDUNG 9.24: Groovy-Skript für die Berechnung der Kontextinformation „roomAirQuality“.

In Abbildung 9.24 ist ein Beispiel für ein Groovy-Skript dargestellt, das für die Interpretation der Luftqualität genutzt wurde. Es suchte zunächst in der Datenbank nach passenden Sensoren, in diesem Fall dem VOCs-Sensor, der im Sensormodell als Sensor vom Typ „airQuality“ mittels der Metadaten abgebildet wurde. Falls (noch) keine Sensoren oder Daten vorhanden waren, wurde die Berechnung vorerst übersprungen. Für jeden Sensor des Typs „airQuality“ wurden daraufhin Daten über die letzten 30 Minuten aus der Datenbank abgefragt, der Mittelwert berechnet und mit Schwellwerten bestimmt, ob die Luftqualität „gut“, „neutral“ oder „schlecht“ war. Die Schwellwerte orientierten sich in diesem Fall an denen des Sensorherstellers. Das Ergebnis wurde mit der Information über den Raum, in dem der Sensor platziert war, und die genutzte Sensorik im Speicher abgelegt.

Das Empfehlungssystem, das den aktuellen Kontext benötigte, wurde aktiviert, sobald die Berechnung der Nutzerpräsenz oder Raumpräsenz durchgeführt wurde. Es konnten jedoch auch andere Kontextinformationen hierfür genutzt werden. Das Empfehlungssystem erhielt den Zeitstempel mit dem aktuellsten berechneten Kontext über den Eventbus und konnte daher die Daten auslesen, ohne den Dateninterpreter zu blockieren.

Gegenüber dem ersten Demonstrator wurde die Erfassung der Nutzerpräsenz deutlich stabilisiert, indem nun die Daten der beiden Abstandssensoren (IR und Ultraschall) sowie ein eventueller Tastendruck berücksichtigt wurden. Sobald eine Sensorquelle eine Präsenz registrierte, wurde die Nutzerpräsenz vor dem Tablet erkannt. Die Raumpräsenz wurde mittels des Bewegungsmelders bestimmt und war für den ambienten Modus des Systems relevant. Die physische Nutzeraktivität wurde abgeschätzt, indem mittels des Stromverbrauchs des TV-Geräts, dessen Laufzeit bestimmt wurde. Es wurde davon ausgegangen, dass der Nutzer während der Laufzeit vor dem TV-Gerät saß und somit sich nicht ausreichend bewegte.

Das Empfehlungssystem nutzte bei diesem Demonstrator zusätzlich die Bewertungen der Empfehlungen sowie die Nutzerstimmung bei der Empfehlungsauswahl. Es wurden zudem immer mehrere passende Empfehlungen aus jeweils einer Kategorie ausgewählt und vom System dargestellt, durch die zudem navigiert werden konnte. Es wurden hierbei maximal fünf Empfehlungen ausgewählt. Falls neue Postkarten empfangen wurden, sind diese vor den Empfehlungen angezeigt worden. Dem Nutzer war es nun auch möglich, selbst eine Kategorie für Übungen über das Menü auszuwählen, hierbei wurden nicht zum Kontext passende Empfehlungen aussortiert.

Systeminteraktion und Inhalte In Abbildung 9.25 ist die Auswahl der Übungskategorien dargestellt, die über das Menü erreicht werden konnte. Über das Menü war es zudem möglich, alte Postkarten anzusehen und eine kurze Beschreibung der Tastenfunktionen als Hilfe zu erhalten.



ABBILDUNG 9.25: Menü zur manuellen Auswahl von Übungen beim zweiten Prototyp.

Bei diesem Prototyp konnte der Nutzer durch die ausgewählten Empfehlungen mittels Hardware- und Software-Buttons navigieren und zusätzlich eine Pausierung auslösen. Die Pausierung besaß einen Timeout von 5 Minuten, damit diese nicht vergessen wurde. In dieser Zeit wurde die Nutzerpräsenz nicht zurückgesetzt und die aktuelle Ansicht blieb erhalten. Die Pause konnte durch einen weiteren Tastendruck beendet und somit beispielsweise eine Übung fortgesetzt werden. Für Übungen und Aufgaben wurde eine „Aufgabe erledigt“-Taste eingeführt, da diese nicht alle automatisch überwacht werden konnten. Des Weiteren wurden interaktive Übungen eingeführt, bei denen der Touchscreen verwendet werden konnte. Die Überwachung von Übungen mittels spezieller, externer Sensorik, wie einer Bodensensormatte, wurde vorbereitet.

Am bedeutsamsten war, dass das Design aller Empfehlungen überarbeitet worden war. Zudem waren mehr Animationen enthalten, die mittels CSS implementiert werden konnten. Die Anzahl der Empfehlungen wurde mehr als verdoppelt (ca. 130). Dies war insbesondere relevant, da einige Empfehlungen aufgrund der aktuellen Jahreszeit oder Tageszeit herausgefiltert werden konnten.

Ambienter Modus Der ambiente Modus des CARE-Systems wurde um eine RGB-LED-Lampe erweitert. Die Lampe stammte aus dem Milight-Funksystem. Abbildung 9.26 gibt einen Überblick über die Zustände des Prototyps. Nachfolgend werden die Zustände und Übergänge beschrieben.

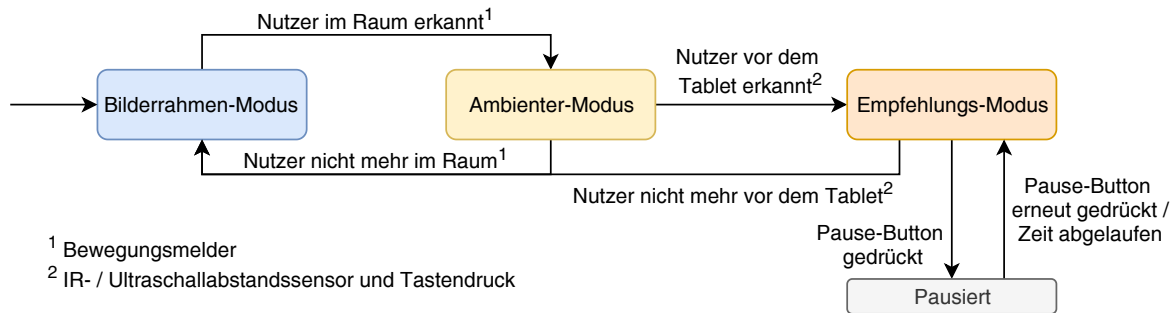


ABBILDUNG 9.26: Überblick über die System-Zustände beim CARE-Prototyp 1.5 / 2. Die genutzten Sensoren sind angegeben. Die Erkennung eines Nutzers im Raum sowie vor dem Tablet waren mit Verzögerungen versehen, bevor ein Zustandsübergang stattfand.

Vor dem Betreten des Raumes war das System im Bilderrahmen-Modus, indem es ausschließlich Bilder anzeigte. Sobald der Raum betreten wurde, wurde das ambiente Licht zusätzlich aktiviert (ambienter Modus). Dieses besaß verschiedene Farben, die eine Vorschau auf die Empfehlungskategorie ermöglichten, zu der passende Empfehlungen am Display angesehen werden konnten. Falls die Farbe „Weiß“ verwendet wurde, wurden nur Postkarten vom System ausgewählt, „Gelb“ entsprach einer Abfrage der Nutzerstimmung, ansonsten wurde die Farbe der Kategorie genutzt, aus der Empfehlungen ausgewählt wurden.

Sobald der Nutzer vor dem System stand, wurde in den Empfehlungs-Modus geschaltet. Eine Pausierung war jedoch möglich, in der der Nutzer sich vom Tablet entfernen konnte, ohne, dass der Empfehlungs-Modus beendet wurde. Diese Pausierung konnte entweder vom Nutzer durch erneutes Drücken der Pause-Taste oder automatisch vom System nach fünf Minuten beendet werden. Sobald der Nutzer nicht mehr vor dem Tablet und im Raum erfasst wurde, erfolgte eine Rückschaltung in den Bilderrahmen-Modus.

9.5.3 Evaluation

Ablauf Die Evaluation des zweiten Demonstrators fand im selben Haushalt statt, wie beim ersten Demonstrator. Diesmal wurde allerdings primär die Seniorin gebeten, das System zu nutzen, da die Personalisierung erweitert wurde. Die Seniorin wurde ausgewählt, da einige der Inhalte für sie bereits beim ersten Demonstrator explizit angepasst wurden und daher ein besserer Vergleich möglich war. Die Studie fand ungefähr ein Jahr nach der ersten statt. Da die Darstellungen der Inhalte komplett überarbeitet wurden und mehr als die doppelte Menge an Empfehlungen sowie neue Funktionalitäten ergänzt worden waren, war dieses System für die Teilnehmerin weitestgehend neu. Die Studie wurde wiederum 14 Tage durchgeführt, wobei das System zwischenzeitlich dreimal von einem Projektmitglied gewartet wurde. Hierbei wurden teilweise kleinere Probleme behoben, die Nutzerin kurz zum System befragt sowie die Daten gesichert.

Um in der Wohnung der Seniorin eine Internetverbindung zu ermöglichen, musste eine UMTS-Verbindung (Datenvolumen: 1 GiB für einen Monat) aufgebaut werden. Durch die Datenbegrenzung musste die Windows-Konfiguration bezüglich automatischer Updates angepasst werden. Das System benötigte den Internetzugang primär zum Empfang der Postkarten und Nutzung lokaler Wetterdaten. Da die Postkarten Bilder enthalten konnten, wurden diese zuvor komprimiert, um das UMTS-Datenvolumen zu schonen. Die Software wurde so entwickelt, dass sie notfalls auch ohne Internetverbindung weiterarbeiten konnte.

Ergebnisse Allgemein gefielen der Seniorin die überarbeiteten und neuen Empfehlungen des zweiten Demonstrators im Vergleich zum ersten Prototypen wesentlich besser.

Es wurden insgesamt 88 Bewertungen von der Seniorin abgegeben, wobei diese sehr ungleich verteilt waren: 86 Bewertungen (74 positiv, 12 negativ) für Postkarten und nur zwei Bewertungen (beide positiv) für Empfehlungen. Auf Nachfrage nach der Evaluation, warum sie die Empfehlungen selten bewertet hatte, antwortete die Seniorin, dass sie die Übungen alle gut fand und diese daher nicht mehr bewertete. Durch die wenigen Empfehlungsbewertungen konnte allerdings das Empfehlungssystem diese kaum bei der Auswahl berücksichtigen. Die Bewertungen von Postkarten wurden ausschließlich protokolliert, da diese nicht selektiert wurden.

Generell schien die Interaktion mit dem Touchscreen des Surface-3-Pro-Tablets der Seniorin leichter gefallen zu sein als mit dem Tablet des ersten Prototyps, da die Gesamtanzahl der Bewertungen wesentlich höher war. Ein Indiz hierfür war auch, dass die Hardware-Buttons zwar immer wieder verwendet wurden, allerdings weniger häufig als die visualisierten Buttons, die ein auditives und visuelles Feedback beim Druck auslösten. In Abbildung 9.27 ist die Nutzung dargestellt. Wie erwartet, wurden die Navigations-Buttons am häufigsten verwendet, der „Aufgabe erledigt“-Button (Erledigt) am seltensten. Die Verteilung der Button-Nutzung ist bei den Hardware- (Key) und Software-Buttons (Touch) ähnlich. Der „Erledigt“-Knopf sollte gedrückt werden, sobald eine Übung durchgeführt wurde. Dies geschah allerdings sehr selten, eventuell da die Seniorin keinen Sinn dahinter sah.

Zehn interaktive Rätselaufgaben (Quiz) wurden in der Studienlaufzeit beantwortet. Acht davon korrekt. Die Fragen wurden somit nicht zufällig beantwortet bzw. waren für die Seniorin lösbar. Die Nutzerstimmung wurde von der Teilnehmerin sechsmal angegeben.

Die Möglichkeit, das Menü zu nutzen, um gezielt Kategorien auszuwählen, wurde mehrfach in Anspruch genommen. Am häufigsten wurden laut den aufgezeichneten Daten die Postkarten (40-mal) gezielt ausgewählt, danach folgten die Anzeige eines Bildes (13-mal), Übungen zum physischen (11-mal) und emotionalen Wohlbefinden (1-mal). Hierbei bestätigte sich die Aussage der Seniorin, dass sie vor allem die physischen Übungen bevorzugte. Zudem schien ein hohes Interesse an den Postkarten zu bestehen. Sie bestätigte auch bei einer kurzen Befragung, dass sie die physischen Übungen regelmäßig durchführte und diese daher gezielt über das Menü auswählte und auch die Postkarten explizit ansah.

Die Farbe der aufgebauten ambienten LED-Lampe schien bei einer kurzen Befragung von der Seniorin nicht verstanden worden zu sein und war daher für sie nicht selbsterklärend. In Zukunft wäre es daher sinnvoll, den Bezug der Lampenfarbe hinsichtlich der Farben der Kategorien (Mapping) explizit zu erklären, da diese Funktion grundsätzlich nützlich sein könnte.

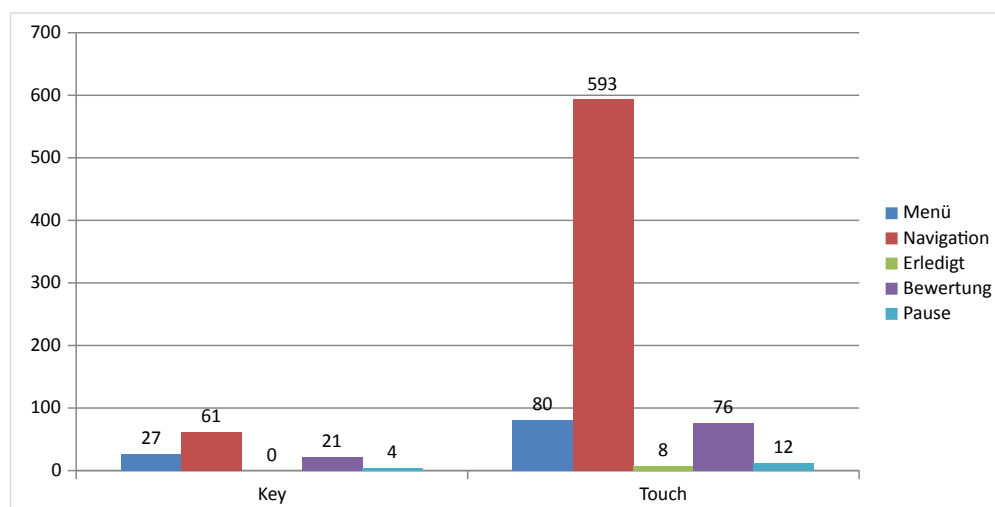


ABBILDUNG 9.27: Anzahl der Buttondrück-Ereignisse der Hardware- (Key) und Software-Buttons (Touch) während der Studienlaufzeit.

9.5.4 Zusammenfassung

Mit dem zweiten Prototyp konnten die wesentlichsten Probleme des ersten Demonstrators behoben werden. Diese betrafen vor allem die Stabilität von Hard- und Software sowie die Interaktion mit dem System.

Durch die Erweiterungen konnte erhöhtes Interesse am Prototyp erzielt werden. Insbesondere durch den Postkartenmodus wurde die Seniorin, die zuvor nicht über einen Internetzugang verfügte und keinen Computer nutzte, besser in ihre Familie eingebunden. Verwandte konnten ihr nun auch Bilder senden und gegebenenfalls per Telefon mit ihr darüber reden, wodurch umfassendere Gespräche möglich waren. Durch die erweiterten Interaktionsmöglichkeiten wurden vor allem die physischen Übungen öfters aufgerufen und konnten nun gezielt bearbeitet werden. Die Übungen ließen sich somit auch besser in den Tagesablauf der Seniorin einbinden.

Allerdings gibt es noch Spielraum für weitere Verbesserungen und Anpassungen des CARE-Prototyp 2. Beispielsweise sollte der „ambiente Modus“ noch gezielter untersucht werden, um dessen Auswirkungen bei diesem System beurteilen zu können. Ohne vorherige Funktionserklärung ist er, generell betrachtet, offensichtlich nicht leicht verständlich und zugänglich.

Die Häufigkeit von Bewertungen sollte gesteigert werden, um die Auswahl der Empfehlungen zu optimieren und deren Akzeptanz beurteilen zu können. So war es nachteilig, dass die Seniorin die Übungen alle gut fand und diese daher nicht bewertete, sodass das System diese Information nicht für die Auswahl der Empfehlungen nutzen konnte. Generell wurden verhältnismäßig wenige Bewertungen abgegeben. Zusätzlich wurde die manuelle Bestätigung mit dem Button „Aufgabe erledigt“ sehr selten verwendet. Damit wurde die Beurteilung erschwert, ob bestimmte Empfehlungen umgesetzt wurden. Dies wäre nur mit einer nahezu vollständigen Überwachung erreichbar gewesen, die jedoch aus Gründen des Schutzes der Privatsphäre unerwünscht war.

Als Konsequenz wurde daher nach Möglichkeiten gesucht, Anreize zu geben, dass Systemnutzer bestimmte Interaktionen häufiger ausführen, da diese beispielsweise für eine personalisierte Auswahl der Empfehlungen benötigt werden. Im nachfolgenden Abschnitt werden Gamifizierungen betrachtet, die hierfür einen Lösungsansatz darstellen können.

9.6 Entwicklung CARE-Prototyp 2 mit Gamifizierungen

Die in diesem Abschnitt beschriebenen Studien wurden in der eigenen Publikation (Herpich, Rist, [Seiderer](#) & André, 2017) veröffentlicht.

9.6.1 Gamifizierung

Der Begriff Gamifizierung oder Gamifikation (engl. gamification) wurde von Deterding et al. (2011) so definiert, dass hierbei Spielelemente in einem nicht spielbezogenem Kontext eingesetzt werden. Ein Beispiel sind verschiedene Arten von virtuellen Belohnungen z. B. in Form von Abzeichen (engl. badges), die beispielsweise für das Erreichen einer bestimmten Schrittzahl in einer Fitness-App angezeigt werden. Für eine Gamifizierung des CARE-Systems waren insbesondere Studien und Projekte von Relevanz, die einen gesünderen Lebensstil und eine Verbesserung des Wohlbefindens fördern. Verheijden Klompstra et al. (2014) analysierten 11 Studien, die den Einfluss von Exergames auf die physische Aktivität von Senioren untersuchten. Die Autoren schlussfolgerten, dass Exergames die physische Aktivität von Senioren, Schlaganfallpatienten und Herzpatienten verbessern können, allerdings weitere Untersuchungen notwendig wären, um Langzeiteffekte beurteilen zu können. 33 Studien über gamifizierte Anwendungen für das Training und die Bestimmung der kognitiven Leistungsfähigkeit wurden von Lumsden et al. (2016) begutachtet. Sie fanden verschiedene Gründe, Spielelemente in Anwendungen zu nutzen, wozu auch eine Steigerung der Langzeitmotivation, Verbesserung der Usability, intuitive Nutzbarkeit und Attraktivität der Anwendung zählten. D. Johnson et al. (2016) analysierten 19 Arbeiten bezüglich

der Effekte von Gamifizierungen bei physischen und mentalen Übungen auf die Gesundheit und das Wohlbefinden. Die meisten Studien zeigten hierbei positive Auswirkungen, ca. 41 % stellten gemischte oder neutrale Effekte, insbesondere bezüglich der Kognition, fest.

Die in den Arbeiten untersuchten Studien geben Rückschlüsse darüber, welche Gamifizierungs-Ansätze gut oder weniger gut funktionierten. Es ist allerdings meist unklar, warum dies der Fall war. Bei einer Studie von Zuckerman und Gal-Oz (2014) konnte beispielsweise kein positiver Effekt durch Gamifizierungen festgestellt werden. Die Autoren nahmen hierbei an, dass die Punktevergabe nicht verstanden wurde und somit bedeutungslos für die Nutzer wurde. Es gibt einige Versuche, das Nutzerverhalten bei Gamifizierungen besser zu verstehen, indem die Rolle von Anreizen und Belohnungen von Computerspielen mittels Psychologie diskutiert wurden. Ein Beispiel hierfür war der Ansatz von Richter et al. (2015). Hierbei wurde die Selbstbestimmungstheorie (siehe hierzu Ryan und Deci (2000)) herangezogen, die zwischen intrinsischer und extrinsischer Motivation unterscheidet. Intrinsisch motiviertes Verhalten stammt aus eigenem Interesse, extrinsisch motiviertes Verhalten wird durch äußere Anreize ausgelöst. Richter et al. (2015) stuften z. B. Punkte, Abzeichen, Ranglisten als extrinsische Motivatoren ein. Die Autoren fassten zusammen, dass bei einer sorgfältigen Auswahl und Implementierung extrinsischer Motivatoren, intrinsische Motivation ausgelöst werden kann, die dabei hilft das Verhalten beizubehalten. Allerdings ist nicht garantiert, dass ein in einem bestimmten Kontext gut funktionierender Gamifizierungs-Mechanismus auch in anderen Kontexten funktioniert.

Ein Großteil der Studien legt nahe, dass Gamifizierungen in vielen Anwendungsfällen funktionieren können. Es bleibt allerdings eine Herausforderung zu entscheiden, welche Elemente genutzt oder ausgeschlossen werden sollten, da diese auch gegenteilige Wirkung ausüben können. Vorschläge für das Design können aus Arbeiten, wie beispielsweise von De Schutter und Vanden Abeele (2010), gewonnen werden, die Leitfäden für Gamifizierungen, basierend auf Interviews mit Senioren, erstellten.

9.6.2 Workshops

Um die grundlegenden Anforderungen zur Gestaltung der Gamifizierung für das CARE-System zu erhalten, wurden zwei Workshops mit Senioren durchgeführt.

Hierbei sollten die täglichen Routinen der Teilnehmer festgestellt werden, insbesondere auch im Hinblick darauf, inwieweit sie bereits Maßnahmen für einen gesünderen Lebensstil getroffen hatten. Zudem sollte Feedback über Technologien eingeholt werden, die Empfehlungen geben und Aktivitäten vorschlagen. Für die Gamifizierungen sollten Meinungen für ein sinnvolles Belohnungsschema für die empfohlenen Aktivitäten gesammelt werden. Die Workshops wurden in einem Seniorentreff sowie in einem Seniorenheim durchgeführt, sodass auch der Einfluss der verschiedenen Lebensumgebungen untersucht werden konnte. Im Folgenden werden der Ablauf und die Resultate zusammengefasst.

Ablauf

Für den ersten Workshop wurden 12 Senioren (11 Frauen, 1 Mann) eines lokalen Seniorentreffs im Alter zwischen 60 und 70 Jahren akquiriert. Sie waren alle im Ruhestand und führten ihren eigenen Haushalt. Für den zweiten Workshop wurden neun Bewohner eines Seniorenheims (8 Frauen, 1 Mann) befragt, die über 80 Jahre alt waren. Im Vergleich zu den jüngeren Teilnehmern des ersten Workshops besaßen diese wesentlich größere altersbedingte Einschränkungen, wie beispielsweise reduzierte Beweglichkeit und teilweise verringerte Kognition.

Um Informationen über die täglichen Routinen und Aktivitäten sowie die Haltung der Senioren gegenüber verschiedenen Belohnungsschemata zu erhalten, wurden ein Arbeitsblatt mit Zeitleiste und drei bedruckte Kartenstapel vorbereitet. Ein Stapel bestand aus Karten, in denen Alltagsroutinen dargestellt waren. Ein weiterer Stapel enthielt Karten, die Aktivitäten darstellten, die das

CARE-System empfehlen würde, um einen gesünderen Lebensstil zu unterstützen. Beim letzten Stapel waren drei verschiedene Belohnungen enthalten, die für CARE geeignet erschienen. Hierzu zählten Punkte, Abzeichen und virtuelle Währungen. Zudem konnten die Teilnehmer leere Karten nutzen, um eigene Vorschläge zu ergänzen.

Zusätzlich wurden semistrukturierte Interviews durchgeführt. Hierbei wurden Fragen gestellt, wie beispielsweise, ob die Teilnehmer bereits einen gesunden Lebensstil führten und wie sich dies in deren täglichen Routinen widerspiegelte oder welche Arten von Belohnungsschemata den Teilnehmern bekannt waren oder was sie über ein Empfehlungssystem wie CARE denken.

Zusätzlich wurden zwei physische Übungen des CARE-Systems (Prototyp 2) vorgestellt, die von den Teilnehmern durchgeführt werden konnten und auf Übungen basierten, die speziell für Senioren angedacht waren. Bei der ersten Übung sollten im Sitzen leichte Hanteln, mit gestreckten

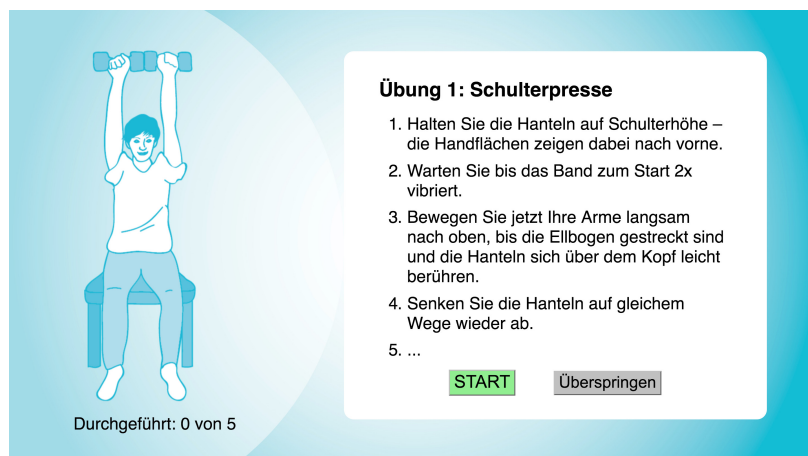


ABBILDUNG 9.28: Darstellung der mit einem Microsoft Band 2 und GRT überwachten Übung.

Armen, hochgehoben werden (siehe Abbildung 9.28). Die Instruktionen waren in textueller Form und als Animation sichtbar. Der Fortschritt der Übung wurde in Echtzeit mittels der Accelerometer-Daten eines Microsoft Band 2 überwacht. Die Signalverarbeitung basierte auf dem „Gesture Recognition Toolkit“ (GRT) von Gillian und Paradiso (2014), das eine DTW-Implementierung einsetzte.

Bei der zweiten Übung sollte das Halten des Gleichgewichts trainiert werden. Die Visualisierung während der Übung ist in Abbildung 9.29 dargestellt. Die Übung wurde ebenfalls textuell und per Animation dargestellt. Zusätzlich wurde der aktuell erkannte Zustand angezeigt, der mittels einer Tanzmatte und SSI ausgewertet wurde. Ein Zustandsautomat war hierbei innerhalb einer SVG-Vektorgrafik als Javascript-Code enthalten, der per WebSocket-Verbindung von SSI die ausgewerteten Werte der Matte empfing.

Resultate

Die täglichen Routinen unterschieden sich zwischen den Teilnehmern der beiden Workshops deutlich, da die Senioren im Seniorenheim beispielsweise keinen Haushalt führen mussten und ihnen somit mehr Zeit für Freizeitaktivitäten zur Verfügung stand. Zusätzlich wurden im Seniorenheim häufig von den Mitarbeitern Aktivitäten für den Nachmittag eingeplant.

Bezüglich des Zeitpunkts von Empfehlungen wurden von den Teilnehmern des ersten Workshops für den Morgen insbesondere die Empfehlungen zum Lüften und Sport treiben ausgewählt. Für den Rest des Tages wurden entspannende Empfehlungen bevorzugt. Einige der Empfehlungen dienten offenbar zur Erinnerung an Aktivitäten, die bereits im Alltag verankert waren, wie beispielsweise Sport. Andere Empfehlungen wurden augenscheinlich zur Stimulation ausgewählt, um zum Beispiel Entspannungsübungen neu in den Alltag zu integrieren, da diese zuvor nicht in

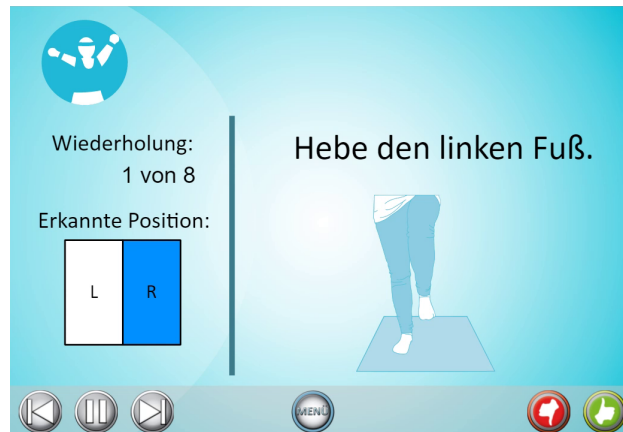


ABBILDUNG 9.29: Darstellung während einer Übung mit der Tanzmatte. Hier wurde gerade der linke Fuß hochgehoben (weiß), sodass der nächste Zustand nach einer Verzögerung aktiviert wurde. Da die Übung automatisch erfasst werden konnte, war der „Aufgabe erledigt“-Button ausgeblendet.

den individuellen Übersichten über die Alltagsroutinen auftraten. Die Teilnehmer des Workshops im Seniorenheim bevorzugten ebenfalls Empfehlungen zum Lüften und Sport am Morgen. Am Nachmittag war dies jedoch sehr personenabhängig. Einige der Teilnehmer wünschten sich auch Empfehlungen zur Erinnerung an geplante Gruppenaktivitäten wie Singen oder Kartenspielen zu erhalten. Empfehlungen bezüglich sozialen Aktivitäten wurden von Teilnehmern des ersten Workshops nicht ausgewählt.

Bei den ihnen bekannten Belohnungsschemata bezogen sich einige der Senioren auf ihre Schulzeit, wo sie Sportabzeichen, Sticker oder Stempel erhielten. Die meisten kannten jedoch auch kommerzielle Systeme wie zum Beispiel Payback. Für die verschiedenen Teilnehmer gab es allerdings auch sehr unterschiedliche Arten der Belohnung, wie beispielsweise Lob von Freunden. In vielen Fällen wurde allerdings die Aktivität selbst oder das Ergebnis von Aktivitäten bereits als belohnend erachtet. Nur vier der neun Teilnehmer aus dem Seniorenheim wollten überhaupt Belohnungen erhalten. Diese wurden in Form virtueller Währung und Abzeichen bevorzugt. Bis auf eine Ausnahme waren alle Teilnehmer des Seniorentreffs offen für Belohnungen. Sie bevorzugten virtuelle Währungen deutlich. Den Aktivitäten wurden bei beiden Workshops verschiedene Belohnungen zugewiesen und waren sehr individuell.

Die physischen Übungen, die von den Teilnehmern ausprobiert werden konnten, wurden von acht Senioren des Seniorentreffs (7 Frauen, 1 Mann) durchgeführt. Bei der Armübung mit dem Armband wurde die animierte Anleitung bevorzugt, da sie sich mit den Bewegungen synchron bewegen konnten. Das Microsoft Band 2 wurde nicht als störend empfunden, allerdings konnten nicht alle Teilnehmer, ohne Hilfe durch die Versuchsleiter, das Band ausreichend straff anziehen. Die Gleichgewichtsübung mit der Tanzmatte unterforderte die meisten dieser jüngeren Senioren. Für einige der Senioren im Seniorenheim waren die Übungen allerdings zu schwer, da hier häufiger stärkere körperliche Einschränkungen vorhanden waren. Der Schwierigkeitsgrad der Übungen müsste angepasst und sie je nach Einschränkungen und Krankheiten selektiert werden, da ungeeignete Übungen im schlimmsten Fall zu Verletzungen führen könnten.

9.6.3 Umsetzung im CARE-System

Aus den Resultaten der Workshops wurden einige Ideen für Gamifizierungen aufgegriffen, um diese im CARE-System praktisch umzusetzen. Hierbei wurde insbesondere berücksichtigt, dass für die Verbesserung des Empfehlungssystems notwendige Systeminteraktionen, wie die Bewertung von Empfehlungen und die manuelle Kennzeichnung erfüllter Vorschläge („Aufgabe erledigt“), gefördert wurden. Vom System überwachbare Empfehlungen, wie beispielsweise Rätselaufgaben

oder interaktive Übungen, konnten bereits automatisch als erledigt gekennzeichnet werden. In den Workshops zeigte sich, dass virtuelle Währungen am beliebtesten waren. Es wurden bei der Implementierung daher Punkte als Währung gegeben, die in eine physische Belohnung, z. B. Eisessen, umgetauscht werden konnten. Eine Übersicht über bisher verdiente Punkte sollte den Teilnehmern ihren Fortschritt verdeutlichen und die weitere Umsetzung der Empfehlungen fördern (siehe Abbildung 9.30 oben rechts).

Die Übersichten waren über das Menü erreichbar (siehe Abbildung 9.30 oben links). Übungen gaben, je nach Umfang, die meisten Punkte, reine Bewertungen nur einen Punkt. Eine Ausgabe mittels einer für einige Sekunden angezeigten grünen Pop-Up-Anzeige für die Bepunktung einer erledigten Aufgabe bzw. Bewertung ist in Abbildung 9.30 unten links bzw. rechts dargestellt. Die Erfüllung der meisten Empfehlungen musste vom Nutzer manuell bestätigt werden, sodass der Nutzer zu sich selbst ehrlich sein sollte, was bei der freiwilligen Nutzung eines Empfehlungssystems wie CARE allerdings eine Grundvoraussetzung ist, um positive Effekte auf die Gesundheit und das Wohlbefinden zu ermöglichen.



ABBILDUNG 9.30: Screenshots vom CARE-System mit implementierten Gamification-Elementen.

9.6.4 Evaluation

Ablauf Mit den oben vorgestellten Anpassungen wurde das CARE-System (Prototyp 2) erneut in der Wohnung (Küche) der Studienteilnehmerin in Augsburg installiert. Der Aufbau des Systems war nahezu identisch zu dem der vorherigen Studie, um Verhaltensveränderungen der Person aufgrund des adaptierten Systems möglichst exakt nachvollziehen zu können. Der Studienzeitraum umfasste diesmal den Zeitraum zwischen Anfang August bis Anfang September (ca. fünf Wochen). Es wurde die ambiente Lampe entfernt, da deren Auswirkungen nicht weiter untersucht werden sollten. Das System hatte in der Zwischenzeit einige Optimierungen und Korrekturen erfahren, die die Stabilität verbesserten. Es wurde wiederum ein UMTS-Datenpaket mit 1 GiB Volumen für

einen Monat gebucht, sodass aktuelle Wetterdaten und der Postkartenmodus verfügbar waren. Für die Auswertungen wurden der erste und letzte Tag ausgeschlossen, da an diesen der Aufbau / Test bzw. Abbau des Systems stattfanden.

Ergebnisse Es wurden während dieses Zeitraums insgesamt 324 Empfehlungen angezeigt, wobei es sich um 54 unterschiedliche handelte. Es wurden insgesamt 34 Bewertungen für Empfehlungen abgegeben, die alle positiv waren. Allerdings wurde keine Aufgabe als erledigt gekennzeichnet und somit nur Punkte für die Bewertungen und das Beantworten der Nutzerstimmung vergeben. Im Studienzeitraum wurden insgesamt sechs Postkarten von Verwandten empfangen.

Zusätzlich wurde die Nutzung der Touchscreen- und Hardware-Buttons ausgewertet, um einen Vergleich mit der vorherigen Studie zu ermöglichen. Insgesamt wurden die Taster 808-mal (ca. 81 %), der Touchscreen 189-mal (19 %) verwendet. Der Großteil der Button-Interaktionen wurden bei beiden Studien durch die Navigation verursacht, hierbei wurden die Taster zu ca. 84 %, bei der vorherigen Studie zu 9 % eingesetzt. Generell wurden diesmal allerdings die Tasten (ca. 81 % vorherige Studie ca. 13 %) bevorzugt. Dies galt jedoch nicht für die Bewertungen (0 %; vorherige Studie: ca. 22 %).

Die Auswertung der Evaluation. ergab, dass der aktuelle Punktestand über das Menü neunmal abgerufen wurde. Übungen wurden über das Menü 23-mal manuell aufgerufen. Die Anzahl der Bewertungen für Empfehlungen konnte beim gamifizierten CARE-Prototyp 2 deutlich gesteigert werden. Beim vorherigen Prototyp wurden in zwei Wochen zwei Empfehlungsbewertungen abgegeben, beim neuen Prototyp 34 Bewertungen in fünf Wochen.

Da die geringere Anzahl an Postkarten selten bewertet wurde, war die Gesamtanzahl der Bewertungen geringer. Der aktuelle Punktestand wurde über das Menü mehrfach abgerufen. Bei den direkten Systemeingaben scheint es keinen Favoriten zu geben, da bei dieser Studie die Taster bevorzugt wurden, was zuvor nicht der Fall war.

9.6.5 Zusammenfassung

Die Gamifizierungen waren potenziell geeignet, die Anzahl der Bewertungen zu erhöhen. Die Feldstudie konnte hierzu nur erste Anhaltspunkte liefern, daher wären weitere Untersuchungen durchaus sinnvoll. Das eigenständige Abfragen des Punktestandes durch die Studienteilnehmerin deutete darauf hin, dass ein grundsätzliches Interesse an den Belohnungspunkten bestand, wodurch sich auch die Anzahl der (bepunkteten) Bewertungen bei Empfehlungen erhöht hatte. Unklar war allerdings, warum die Aufgaben nicht als „erledigt“ gekennzeichnet wurden, was zu wesentlich mehr Punkten geführt hätte. Wie sich bereits bei den Workshops gezeigt hatte, gab es allerdings auch sehr starke individuelle Präferenzen gegenüber der Gestaltung der Gamifizierung. Die Auswahl der am besten für eine Person geeigneten Art der Gamifizierung ist allerdings nicht trivial und es müssten hierfür erst entsprechende Tests entwickelt werden. Hierdurch wäre es beispielsweise bei zukünftigen Systemen möglich, personalisierte Gamifizierungs-Mechanismen einzusetzen und somit die Attraktivität der Systeme zu steigern.

9.7 Roboterassistent für Senioren

9.7.1 Einleitung

In einer weiterführenden Untersuchung, die in den eigenen Arbeiten (Ritschel, Janowski, [Seiderer](#), Wagner & André, 2019; Ritschel, [Seiderer](#), Janowski, Wagner & André, 2019) veröffentlicht wurde, sind Bestandteile des CARE-Systems verwendet worden, um einen Gesundheitsassistenten für Senioren auf der Basis eines Roboters zu entwickeln. Der Fokus lag hierbei jedoch auch auf einer personalisierten Anpassung der Formulierungen für die Sprachausgabe des Roboters sowie einer

hohen Usability. Als Inhalte wurden neben Spielen, Empfang von Chat-Nachrichten, Wettervorhersagen, Nachrichten und Terminerinnerungen auch ausgewählte Empfehlungen des CARE-Systems verwendet. Es wurde in diesem Fall allerdings kein kontextsensitives Empfehlungssystem genutzt. Inspiriert durch das CARE-System wurde ein Bedienpult zur Steuerung des Systems entwickelt, um den Senioren eine leichtere Bedienung und das explizite Feedback in Form von Bewertungen zu ermöglichen. Die Inhalte wurden vom System mittels eines Reeti-Roboters und Monitor per Bild, Text und Sprache ausgegeben. Alle Nutzereingaben fanden mithilfe eines Bedienpults statt. In Abbildung 9.31 ist das System skizziert dargestellt.

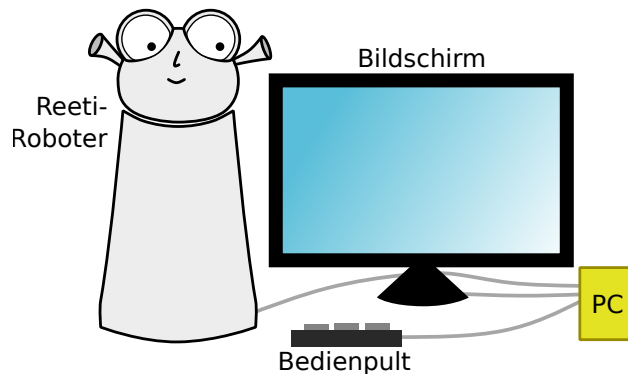


ABBILDUNG 9.31: Schematische Darstellung des Roboterassistenten für Senioren.

In den letzten Jahren sind verschiedene assistierende Roboter für den Heimgebrauch auf den Markt gekommen, die beispielsweise bei der Suche nach Informationen helfen oder das emotionale und mentale Wohlbefinden verbessern sollen. Diese Roboter besitzen in den meisten Fällen entweder ein mechatronisches Gesicht (z. B. Mabu¹⁵ und ElliQ¹⁶) oder es wird ein Gesicht auf einem Bildschirm dargestellt (z. B. Jibo¹⁷ und Buddy¹⁸). Die Interaktion mit dem Nutzer wird mittels Sprachinterface, mobiler Anwendung oder Berührung ermöglicht. Ein Beispiel für einen gesundheitsbezogenen Roboterassistenten mit Display, der die Einnahme von Medikamenten überwachen und diese auch ausgeben konnte, war Pillo von Tao et al. (2016). Dieser konnte zusätzlich Nutzer an ihren Gesundheitsplan erinnern und beantwortete Fragen bezüglich der Inhaltsstoffe von Nahrungsmitteln.

Neben kommerziellen Produkten werden auch in der Forschung Assistenzroboter eingesetzt. Hierzu gehörte der soziale Roboter, der im Rahmen des AlwaysOn-Projekts entwickelt wurde (Sidner et al., 2018). Hierbei wurden ein Reeti-Roboter und zusätzlich ein Bildschirm eingesetzt. Das System ermöglichte verschiedene Arten von Unterhaltungsaktivitäten, wie beispielsweise Spiele, Kommunikation mit anderen oder Übungen für die Verbesserung des Wohlbefindens des Nutzers. Erwähnenswert ist auch die Möglichkeit des Systems, eine Beziehung zum Nutzer aufzubauen, indem Aktivitäten geplant, aber sensible Themen wie Gesundheitsempfehlungen am Anfang vermieden wurden. Wie der Name bereits sagt, war das System immer aktiv, sodass Interaktionen jederzeit möglich waren.

Schroeter et al. (2013) präsentierten im Rahmen des CompanionAble-Projekts einen sozialen Assistenzroboter für Senioren, die unter leichten kognitiven Einschränkungen leiden. Der mobile Roboter war mit Smarthome-Technologien verbunden. Er konnte an Termine erinnern, Videoanrufe durchführen, persönliche Gegenstände aufräumen und ein Spiel zur Stimulation der Kognition spielen.

Im Bereich gesundheitsbezogener Roboteranwendungen gewinnt durch Roboter assistiertes Training eine zunehmende Bedeutung (Tsiakas, Abujelala et al., 2018; Tsiakas, Kyrarini et al., 2018).

¹⁵<http://mymabu.com/> (13.01.2020)

¹⁶<https://elliq.com/> (13.01.2020)

¹⁷<https://www.jibo.com/> (13.01.2020)

¹⁸<https://buddytherobot.com/en/buddy-the-emotional-robot/> (13.01.2020)

Häufig wurden hierbei Gamifizierungen eingesetzt, um den Spaßfaktor und somit auch die intrinsische Motivation zu erhöhen. Schneider et al. (2017) nutzten beispielsweise einen NAO-Roboter¹⁹, um Anweisungen für isometrische Übungen zu geben. Die Ergebnisse zeigten hierbei, dass die Übungen zusammen mit einem humanoiden Roboter effektiver waren, als wenn diese alleine durchgeführt wurden. Roboter wurden auch für Übungen zur Rehabilitation nach einem Schlaganfall (Tapus et al., 2008), Übungen für Demenzzranke (Tapus et al., 2009) oder zur Behandlung autistischer Störungen Liu et al. (2008) eingesetzt. Von Fasola und Mataric (2013) wurden, speziell für Senioren, auch robotergestützte Übungen genauer erforscht, die zur Verbesserung des physischen und mentalen Wohlbefindens beitragen können. Die Roboter gaben hierbei Anweisungen, demonstrierten Bewegungsabfolgen oder Haltungen und gaben Rückmeldung über die Leistung des Nutzers. Die von den Autoren vorgestellten Studienergebnisse zeigten eine eindeutige Präferenz von Robotern gegenüber virtuellen Agenten.

Aktuell für den Heimgebrauch verfügbare Roboter besitzen allerdings häufig nicht entsprechende Aktuatorik und Sensorik, um komplexe Übungen zu demonstrieren und zu überwachen. Daher werden häufig Bildschirme in die Roboter integriert oder zusätzlich zu diesen genutzt, auf denen die Übungen visuell dargestellt werden können. Bildschirme stellen hierbei einen akzeptablen Kompromiss dar. Eine Überwachung der Übungen ist allerdings meist nur rudimentär möglich.

Beim CARE-System waren bereits seit Realisierung der ersten Prototypen visuelle Übungen in Form von Text und Animationen auf einem Monitor (Tablet) verfügbar. Zu einer weiteren Verbesserung der Akzeptanz von Übungen könnte, entsprechend den Forschungsergebnissen, ein geeigneter Roboter durchaus beitragen.

In einer Vielzahl an Experimenten mit Robotern wurde in der Literatur die Sprache des Roboters verändert und untersucht. Durch die Sprache, Gestik, das Blickverhalten und Gesichtsausdrücke kann der Roboter eine bestimmte Persona bzw. Rolle einnehmen.

Bartl et al. (2016) untersuchten bei sozialen Assistenzrobotern für Senioren die Auswirkungen verschiedener Persona auf die Nutzerakzeptanz. Die Persona konnte hierbei in Richtung „Kamerad“ (engl. companion) gehen, wobei der Roboter hierbei eher emotional, enthusiastisch und expressiv war oder in die Richtung „Assistent“, bei der autoritärere und formale Sprache genutzt wurde. Es wurden hierbei die Sprache, Gesichtsausdrücke und Kopfposition angepasst. Ergebnisse der Wizard-of-Oz-Studie zeigten, dass die Persona „Kamerad“ bevorzugt wurde und die Sympathie der Roboter sowie deren wahrgenommene Intelligenz erhöht wurden.

In der Arbeit von Gebhard et al., 2008 konnte bei einem virtuellen Pokerspiel, der virtuelle Roboter die Persona „Mentor“ (kollaborativ) oder „Gegner“ (kompetitiv) annehmen. Beim eigenen System sollte daher bei Spielen ebenfalls untersucht werden, welche Rolle des Roboters bevorzugt werden würde. Die Rollen unterschieden sich hinsichtlich ihrer Persönlichkeit. Die Rolle des „Mentors“ besaß im Gegensatz zum „Gegner“ eine starke Ausprägung der „Verträglichkeit“ und eine schwache Ausprägung des „Neurotizismus“.

Die Ergebnisse der Studie von Huijnen et al. (2011) zeigten, dass neben der Funktionalität, der Charakter und der Interaktionsstil des Roboters wesentlich wichtiger sind als sein Design. Entsprechend unterschiedlicher Präferenzen der Nutzer, sollte der Charakter des Roboters hieran angepasst werden.

Die Bedeutung von Sprachstil und Satzformulierungen eines Assistenzroboters wurde von Hammer et al. (2016) im Hinblick auf deren Einfluss bei Senioren untersucht. Hierbei ergab sich, dass die Wahrnehmung der Höflichkeit stark von der Altersgruppe abhängt und höfliche Formulierungen bei Empfehlungen nicht immer überzeugender wirken. Aus diesen Gründen sollte die Nutzung von höflichen Ausdrücken an den jeweiligen Nutzer angepasst werden.

Unter Berücksichtigung der verwandten Literatur wurde in der eigenen Arbeit ein sozialer Assistenzroboter entwickelt, bei dem der Sprachstil mittels explizitem Nutzerfeedback unter Verwendung von Reinforcement Learning angepasst wurde. Typische Assistenzfunktionen wurden

¹⁹<https://www.softbankrobotics.com/emea/en/nao> (13.01.2020)

implementiert. Hierzu zählten die Informationsabfrage in Form von Nachrichten, Wettervorhersagen und Erinnerung an Termine. Zusätzlich konnten gesundheitsbezogene Empfehlungen (basierend auf dem CARE-System) sowie E-Mails und Chat-Nachrichten ausgegeben werden. Zur Unterhaltung waren ähnliche Spiele wie „Memory“, „Uno“, und „Mensch ärgere Dich nicht“ sowie Witze enthalten. Der Nutzer konnte mittels eines Bedienpults positive oder negative Bewertungen abgeben, die sich auf die Sprache des Roboters bezogen. Anhand dieser wurden mittels Reinforcement Learning die sprachlichen Präferenzen des individuellen Nutzers gelernt und angepasst. Bei der Informationsabfrage wurde die Sprache durch die Auswahl der Persona „Kamerad“ bzw. „Assistent“ (Bartl et al., 2016), bei Spielen durch Einnahme der Rolle „Mentor“ bzw. „Gegner“ (Gebhard et al., 2008) angepasst. Bei den gesundheitsbezogenen Empfehlungen wurde deren Höflichkeit mittels acht Höflichkeitsstrategien von W. L. Johnson et al., 2005 variiert.

9.7.2 Systemaufbau

In Abbildung 9.32 ist der Aufbau des Systems dargestellt. An einem Haupt-PC, der mit dem Internet verbunden war, wurden ein Bedienpult (per USB) und ein Bildschirm angeschlossen. Der Reeti-Roboter verfügte über einen integrierten PC, der per LAN vom Haupt-PC angesteuert wurde, um die Sprachausgaben und Bewegungen zu tätigen.

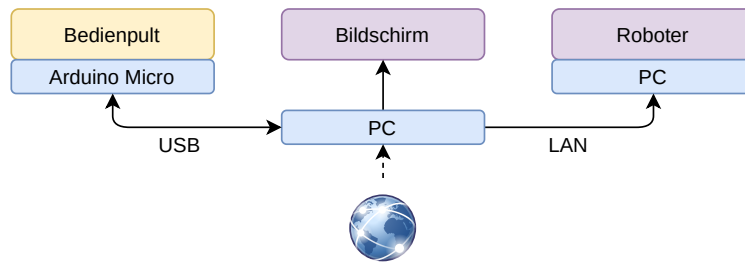


ABBILDUNG 9.32: Übersicht über das System.

Bedienpult Das Bedienpult (siehe Abbildung 9.33) wurde in das System integriert, da bereits beim CARE-System Bedienungsprobleme bei der ausschließlichen Verwendung eines Touchscreens auftraten. Der Aufbau der Elektronik war hierbei sehr ähnlich zu den Hardware-Tastern des CARE-Prototyp 2. Ein Arduino Micro Pro 3,3 V mit 8 MHz (mit ATmega32U4 Mikrocontroller) konnte per USB 2 angeschlossen werden und simulierte eine Tastatur für Tastendruck und war gleichzeitig per serieller Schnittstelle für die Ansteuerung der LEDs ansprechbar. Es wurden 10 mit Einlegern beschriftbare Taster eingesetzt, die jeweils eine weiß leuchtende LED besaßen. Da die Taster und LEDs einzeln ausgelesen werden konnten, wurden insgesamt 20 digitale Pins benötigt, die der Arduino Micro Pro allerdings nicht besaß und zudem die zulässige Maximalstromstärke (200 mA) beim Einschalten aller LEDs überschritten worden wäre. Daher wurde ein I/O-Expander (SX1509) per I²C an den Arduino Micro Pro angeschlossen, der alle LEDs ansteuerte. Die Spannung für die LEDs wurde mittels einer Diode, die allen Anoden vorgeschaltet war, von 3,3 V (vom Spannungsregler des Arduino Micro Pro) auf ca. 2,6 V reduziert, wodurch der zulässige Spannungsbereich der LEDs eingehalten wurde. Die einzelnen Kathoden waren mit den Ports des I/O-Expanders verbunden. Die Taster waren direkt an den Mikrocontroller angeschlossen, wobei die integrierten Pullup-Widerstände per Software eingeschaltet wurden, um undefinierte Zustände zu verhindern und keine zusätzlichen Widerstände in der Schaltung zu benötigen. Die eigens entwickelte Firmware entprellte die Taster, bevor die Tastendrucke an den PC gesendet wurden. Zusätzlich stellte sie die serielle Kommunikation mit binärem Protokoll zur Ansteuerung der einzelnen LEDs bereit. Der Mikrocontroller und I/O-Expander wurden auf Sockelleisten eingesteckt. Um die Schaltung zu schützen und eine Zugentlastung an der empfindlichen Micro-USB-Buchse zu ermöglichen,

wurde ein dreiteiliges, geschraubtes Gehäuse (Unterteil, Abdeckung für Taster, Abdeckung für Mikrocontroller / I/O-Expander) modelliert und mit einem 3D-Drucker gedruckt.

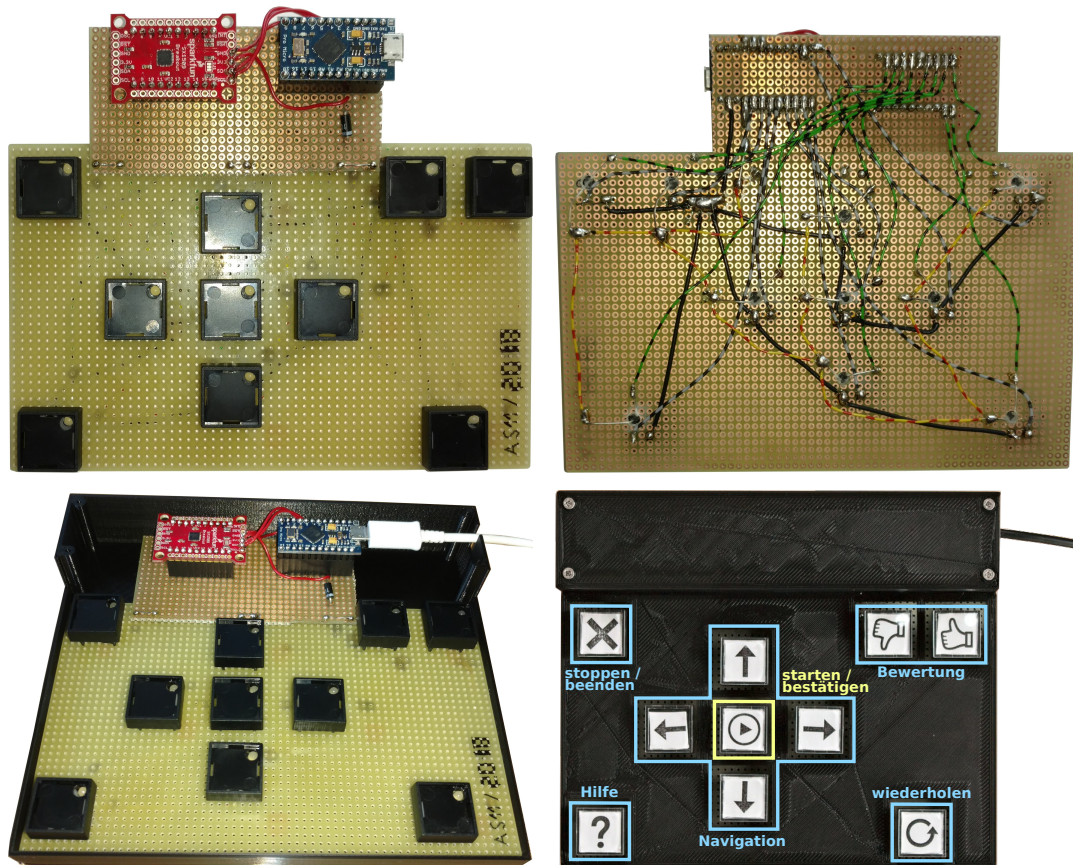


ABBILDUNG 9.33: Bedienpult des Prototyps. Das Pult wurde per USB 2 mit Strom versorgt und als Tastatur erkannt. Die Tasten wurden durch die Firmware entprellt. Die LEDs konnten per serieller Schnittstelle einzeln angesteuert werden. Die Funktion der einzelnen Taster beim System ist unten rechts beschriftet.

Roboteransteuerung Um die Hardware des Roboters vollständig auszunutzen, wurden Animationen verwendet, bei denen das Blickverhalten und Augenblinzeln angepasst wurden. Bei Spielen schaute er beispielsweise auf den Bildschirm oder zum Nutzer und zeigte abhängig vom Spielfortschritt einfache Emotionen. Nach einer Minute Nutzerinaktivität wurde, um die Motoren zu schonen, eine Schlafposition mit geschlossenen Augen eingenommen. LEDs in den Wangen, wurden aktiviert, um anzuzeigen, dass sich das System nur im Standby befand und jederzeit wieder genutzt werden konnte. Der Bildschirm wurde nach drei Minuten abgeschaltet. Durch Eingaben mit dem Bedienpult wurde das System reaktiviert.

9.7.3 Evaluation

Anpassung des Sprachstils Das System wurde in Feldstudien bei zwei deutschsprachigen Senioren in deren Wohnungen für jeweils eine Woche installiert. Es nahmen eine Frau im Alter von 64 Jahren sowie ein Mann mit 61 Jahren an der Studie teil. Die Teilnehmer sollten, so oft sie wollten mit dem System interagieren und besonders auf die Sprache achten. Die Bewertungen des Sprachstils sollten unabhängig vom gesprochenen Inhalt stattfinden. Es wurde den Teilnehmern nicht mitgeteilt, welche Anpassungen bei der Sprache durchgeführt wurden. In Abbildung 9.34 ist die Seniorin mit dem System abgebildet.

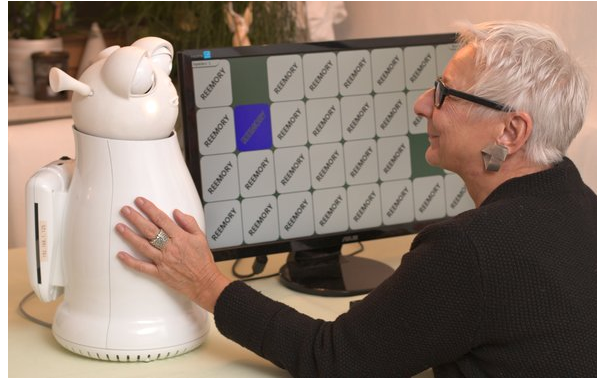


ABBILDUNG 9.34: Studienteilnehmerin mit Prototyp. Das Bedienpult ist hier verdeckt. (Abbildung aus (Ritschel, [Seiderer](#), Janowski, Wagner & André, 2019))

Die Resultate der Studie zeigten, dass die Präferenzen der Teilnehmer tatsächlich sehr unterschiedlich waren, was insbesondere bei der Bewertung der verschiedenen höflichen Formulierungen für die gesundheitsbezogenen Empfehlungen der Fall war. Eine Bevorzugung eines Roboters der Persona „Kamerad“ im Vergleich zum „Assistenten“ konnte, im Vergleich zu anderen Arbeiten, nicht beobachtet werden. Bei der Informationsabfrage trat genau das Gegenteil auf. Während der Studie konnten allerdings zwei Probleme festgestellt werden: Wiederholungen und polarisierende Fremdtex te bei den Informationsabfragen. Wiederholungen könnten zukünftig mittels Textvariationen oder Textgenerierung reduziert werden. Das zweite Problem wäre allerdings schwieriger zu beheben. Polarisierende Nachrichten können nicht vermieden werden und beeinflussen möglicherweise die Bewertungen des Sprachstils.

Generell konnte jedoch festgestellt werden, dass es durchaus möglich ist, dass Nutzer sich nicht darüber bewusst sind, welchen Sprachstil sie bei bestimmten Aufgaben bevorzugen, wodurch ein lernfähiges System vorteilhaft sein kann.

Usability Für das System wurde zusätzlich die Usability mit den beiden Studienteilnehmern untersucht und in (Ritschel, Janowski, [Seiderer](#), Wagner & André, 2019) veröffentlicht.

Um die Technikaffinität zu bestimmen, beantworteten die Studienteilnehmer jeweils einen TA-EG-Fragebogen aus der Arbeit von Karrer et al. (2009). Beide Teilnehmer erreichten hierbei bei der „Kompetenz“ sowie „positiven Einstellung“ relativ hohe Werte. Bei der „Begeisterung“ war die Bewertung der Frau höher als beim Mann, der eher einen neutralen Wert erreichte. Bei der „negativen Einstellung“ lagen beide Teilnehmer unterhalb neutral. Die Ergebnisse zeigten daher eine positive, optimistische Haltung gegenüber elektronischen Geräten und Zuversicht, diese handhaben zu können. Bei dieser Altersgruppe ist dieses Resultat eher ungewöhnlich, dürfte allerdings auch ein Grund für deren Studienteilnahme gewesen sein.

Das Bedienpult wurde mittels des System-Usability-Scale-Fragebogens (Brooke, 1996) von beiden Teilnehmern evaluiert. Es erreichte hierbei sehr hohe Werte 100 bzw. 90 von 100 Punkten.

Die allgemeine Wahrnehmung des Systems wurde mittels des AttrakDiff-Fragebogens von (Hassenzahl et al., 2003) untersucht. Die Werte lagen bei beiden Senioren in allen Kategorien über dem neutralen Wert. Daher wurde von diesen das System hinsichtlich seiner pragmatischen sowie hedonischen Qualität positiv eingestuft.

Bei einer weiteren Befragung erwähnten beide Studienteilnehmer, dass sie mit dem Roboter auch gerne per Spracheingabe interagieren wollen würden. Die Seniorin würde dies allerdings nur akzeptieren, falls der Schutz ihrer Privatsphäre gewährleistet wäre, was bei Cloud-Services nicht der Fall wäre. Zusätzlich hätte sie lange Sprachausgaben gerne früher unterbrochen. Der Senior hatte bemerkt, dass zwischen den Höflichkeitsformen gewechselt wurde, da diese offensichtlich beim Lernen zu auffällig angepasst wurden.

Generell ist festzuhalten, dass der Roboter für die beiden vergleichsweise technikaffinen Senioren problemlos nutzbar war. Insbesondere die Evaluation der Sprache des Roboters ist allerdings problematisch, da hierbei eine effiziente Nutzung relevanter sein könnte als beispielsweise die Höflichkeit.

9.7.4 Zusammenfassung

Bei diesem Projekt wurde ein System entwickelt, das die Grundidee des CARE-Projekts bezüglich einer Verbesserung des Wohlbefindens von Senioren aufgriff und um einen sozialen Roboter erweiterte, um das System attraktiver zu gestalten. Ergänzend wurde ein Bildschirm zur Anzeige von Inhalten eingesetzt. Es handelte sich allerdings nicht um einen Touchscreen. Stattdessen wurde für Systemeingaben ein eigens entwickeltes Bedienpult verwendet, das Ähnlichkeiten zu den Hardware-Tastern des CARE-Prototyp 2 besaß. Vom CARE-Projekt wurden die allgemeinen Empfehlungen zur Verbesserung des Wohlbefindens übernommen. Ebenso ermöglichte das System die Kommunikation mit Verwandten über das Internet. Im Vergleich zum CARE-System waren umfangreichere Spiele zur Unterhaltung enthalten und es konnten Informationen, wie z. B. das Wetter, über das System direkt abgerufen werden. Bei dem für diese Untersuchung entwickelten System standen im Vergleich zum CARE-System die Formulierungen der textuellen Ausgaben des Roboters im Vordergrund, die an den Nutzer angepasst wurden. Hierdurch sollte das System auch längerfristig Interesse wecken und somit attraktiver für Nutzer sein. Die Anpassung erfolgte mittels Reinforcement Learning und basierte auf dem positiven oder negativen Nutzerfeedback, das über das Bedienpult abgegeben wurde. In der Untersuchung konnte der Roboter bei den Spielen, je nach Nutzerpräferenz, die Persona „Mentor“ oder „Gegner“ annehmen, bei der Informationsabfrage „Kamerad“ bzw. „Assistent“. Bei den Empfehlungen wurde die Höflichkeit an den Nutzer angepasst. Die Untersuchungen mit zwei Senioren zeigten, dass die Senioren unterschiedliche Formulierungen bei den verschiedenen Systemanwendungen bevorzugten und es somit von Vorteil sein kann, die Formulierungen eines solchen Systems an den Nutzer anzupassen. Des Weiteren zeigte die Untersuchung bezüglich der Attraktivität und Usability des Systems, dass das System auf die Senioren attraktiv wirkte und u. a. durch den Einsatz des Bedienpults eine hohe Usability besaß.

9.8 Zusammenfassung des Kapitels

In diesem Kapitel wurde zunächst der nutzerzentrierte Entwicklungsablauf des kontextsensitiven, Privatsphäre respektierenden Empfehlungssystems zur Verbesserung des Wohlbefindens von Senioren „CARE“ behandelt. Hierfür wurde ein erster Prototyp entwickelt, dessen Hardware / Software im Laufe der Untersuchung bis zum CARE-Prototyp 2 weiterentwickelt wurde. Die CARE-Prototypen wurden von zwei Senioren, mit geringen technischen Kenntnissen, in Langzeitstudien evaluiert. Die Sensorik wurde so ausgewählt, dass sie in die Wohnumgebung der Senioren leicht integriert werden konnte, hinreichend stabil in ihrer Funktion war und mittels Dateninterpretation relevante Kontextinformationen für das Empfehlungssystem bereitstellen konnte. Im Entwicklungsprozess konnten spezifische Probleme der Zielgruppe Senioren bei der Systeminteraktion deutlich reduziert werden.

Für ein Empfehlungssystem zur Verbesserung des Wohlbefindens ist es von besonderer Bedeutung eine langfristige Nutzung zu erreichen, da anhaltende, positive Effekte durch das Befolgen von Empfehlungen erst nach einer längeren Zeitdauer auftreten können. Eine grundsätzliche Möglichkeit Interesse auf ein System zu lenken besteht darin, Neugierde zu wecken. Beim CARE-System konnte Interesse ausgelöst werden, indem dem Nutzer regelmäßig unbekannte Empfehlungen ausgegeben wurden. Hierfür wurde allerdings eine nicht unerhebliche Anzahl an Empfehlungen benötigt, die während der Entwicklung des CARE-Systems regelmäßig ergänzt wurden. Die Empfehlungen wurden zusätzlich nur in bestimmten, geeigneten Kontexten angezeigt. Des Weiteren

wurde das Interesse am System gesteigert, indem es um einen Postkartenmodus und um Gamifizierungen erweitert wurde. Der Postkartenmodus ermöglichte es, Verwandten, Bilder mit Text an die Senioren über das Internet zu senden. Die Kommunikation über das Internet ist in der Regel älteren Senioren verschlossen, da sie häufig weder über einen (mobilen) Internetanschluss noch über die technischen Fähigkeiten zu dessen Nutzung verfügen. Daher erlaubte die Integration des Postkartenmodus in das CARE-System eine neue Form der Kommunikation mit den Senioren und stieß bei diesen auch auf reges Interesse. Über die Inhalte der Postkarten konnten die Senioren mit ihren Verwandten mittels Telefon sprechen. Insgesamt hatte diese Funktionalität das Potenzial, die Senioren stärker in die Familie einzubinden und eine Vereinsamung zu verhindern. Um das Interesse der an der Studie beteiligten Senioren zusätzlich zu wecken, war eine Belohnungsstrategie mittels Gamifizierung beim CARE-System untersucht worden. Diese Strategie basierte auf den Resultaten von zwei mit mehreren Senioren durchgeführten Workshops. Es wurden Spielelemente in das CARE-System integriert, die in einer Studie unter anderem dafür sorgten, dass mehr Bewertungen für die Empfehlungen abgegeben wurden. Diese Nutzerbewertungen sind für ein Empfehlungssystem essenziell, da hierauf die personalisierte Auswahl der Empfehlungen basiert.

In einer weiterführenden Studie wurden Empfehlungen des CARE-Systems mit einem sozialen Roboter genutzt. In der Studie mit zwei Senioren wurde der Sprachstil durch Reinforcement-Learning angepasst. Der Roboter stellte, neben dem Bildschirm, ein zusätzliches Ausgabemedium dar. Dieser Prototyp profitierte von den Empfehlungen (Inhalte) und Erkenntnissen aus der Entwicklung des CARE-Systems.

Viele der in dieser Arbeit bereits behandelten Technologien wären im Rahmen eines zukünftigen, erweiterten CARE-Systems nutzbar, da diese geeignet sind, weitere Kontextinformationen für ein Empfehlungssystem zur Verfügung zu stellen. Beispielsweise könnten die Nahrungsaufnahme oder die Ausgabemöglichkeiten, z. B. Sprachassistent, erweitert werden. An entsprechenden AAL-Systemen besteht ein immer größer werdendes Interesse. In der aktuellen Literatur kann hierzu das Projekt „NESTORE“ von Angelini et al. (2019) gefunden werden, das gewisse Gemeinsamkeiten mit dem CARE-System besitzt.

Durch die zukünftige Verbreitung von Smarthome-, IoT- und Wearable-Technologie auch in Seniorenhaushalten sowie eine erhöhte Technikaffinität der Senioren wird eine Integration von kontextsensitiven Systemen zur Verbesserung des Wohlbefindens in bereits existierende Infrastruktur zunehmend problemloser realisierbar sein. Einige Grundlagen für die Entwicklung zukünftiger Systeme können bereits unter anderem vom CARE-System und dessen Erweiterungen abgeleitet werden. Diese Systeme sollten besondere Bedürfnisse der Zielgruppe berücksichtigen, allerdings auch so entwickelt werden, dass sie für andere Altersgruppen nützlich sind. Hierdurch könnte ein System ein Leben lang als persönlicher Assistent dienen und neben anderen Aufgaben das Wohlbefinden verbessern. Ubiquitäre Sensorik, Nutzung von Software-Standards sowie Verfahren des maschinellen Lernens, die ein Leben lang lernen, werden hierbei essenziell sein.

Teil IV

Zusammenfassung

10 Zusammenfassung und Ausblick	251
10.1 Wissenschaftliche Beiträge	252
10.2 Fortführende Arbeiten	260
A Publikationen und Reviewaktivitäten	261
A.1 Wissenschaftliche Veröffentlichungen	261
A.2 Durchgeführte Reviews	267
B Lehre	268
B.1 Betreute Lehrveranstaltungen	268
B.2 Betreute Abschlussarbeiten	268
C (mobile)SSI Plugins	269
Literatur	270

10 Zusammenfassung und Ausblick

Das Ziel dieser Arbeit war, Möglichkeiten der Nutzung von Technologien aus dem IoT- und Mobile-Computing-Bereich zur Verbesserung des menschlichen Wohlbefindens zu identifizieren, geeignete Verfahren sowie Assistenz-, Informations- und Empfehlungssysteme zu entwickeln und zu untersuchen. Hierfür wurde ein pragmatisches Modell des menschlichen Wohlbefindens eingesetzt, das im Rahmen des CARE-Projekts entwickelt wurde und fünf Kategorien enthielt. Indikatoren, die eine Abschätzung der verschiedenen Wohlbefindenskategorien ermöglichten, wurden in dieser Arbeit identifiziert und darauf basierende Systeme, die Technologien aus dem IoT- und Mobile-Computing-Bereich nutzten, implementiert. Für die Datenverarbeitung wurden verschiedene Verfahren der Echtzeitsignalverarbeitung und des maschinellen Lernens eingesetzt. Von großer Bedeutung war bei der Umsetzung dieser Systeme eine hohe Nutzerakzeptanz, die bei der Systementwicklung von Beginn an berücksichtigt wurde. Dazu wurden verschiedene Techniken zur Gewährleistung einer hohen Usability und Maßnahmen zum Schutz der Privatsphäre eingesetzt. Des Weiteren wurde bei einigen Systemen die Idee des Ubiquitous Computing, bei dem Systeminteraktionen primär implizit im Aufmerksamkeitshintergrund stattfinden, umgesetzt. Die sich aus der Zielsetzung dieser Arbeit ergebenden Herausforderungen und wissenschaftlichen Forschungsziele wurden definiert, umfassend untersucht und diesbezügliche Lösungen erarbeitet.

Identifikation von relevanten Wohlbefindensfaktoren

Zur Identifikation von Wohlbefindensfaktoren wurde in dieser Arbeit ein Wohlbefindensmodell (siehe Abschnitt 2) eingesetzt, das auf bereits existierenden, etablierten Modellen basiert. Die Kategorien sind hierbei emotionales, mentales, soziales, umgebungsbezogenes und physisches Wohlbefinden. Für diese Kategorien wurden maßgebliche Wohlbefindensfaktoren mittels aktueller Literatur aus Geografie, Psychologie und Humanmedizin identifiziert. Eine sich hieraus ergebende Auswahl geeigneter Faktoren, die im Rahmen dieser Arbeit in Verfahren und Systemen eingesetzt wurden, ist nachfolgend aufgelistet. Sie werden den Wohlbefindenskategorien zugeordnet, wobei einige der Faktoren Einfluss auf mehrere Kategorien besitzen können.

Emotionales, mentales und soziales Wohlbefinden

- **Stress** (siehe Abschnitt 6.1.1) tritt beispielsweise bei einer mentalen Überbeanspruchung auf und besitzt daher einen negativen Einfluss auf das mentale Wohlbefinden. Insbesondere chronischer Stress kann zudem zu gesundheitlichen Problemen führen.
- **Lachen** (siehe Abschnitt 6.1.2) stellt ein nonverbales, soziales Verhalten dar, das in der Regel positive Emotionen ausdrückt. Es kann zudem auf das mentale Wohlbefinden negativ wirkende Stresshormone und negative Gefühle abbauen. Des Weiteren können soziale Bindungen durch Lachen in Gruppen gestärkt werden.
- **Soziale Verbundenheit** (siehe Abschnitt 6.1.3) bezeichnet das Gefühl der Zugehörigkeit und Bindung zwischen Personen und besitzt einen positiven Einfluss auf das soziale Wohlbefinden, da mit ihr positive Gefühle verbunden sind. Ein positiver Effekt auf das mentale Wohlbefinden besteht u. a. dadurch, dass sie Stressabbau unterstützen.
- Die **mentale Fitness** (siehe Abschnitt 9.2.2) kann durch ausgewählte, regelmäßige geistige Übungen verbessert werden. Hierzu zählen z. B. das Training der Merkfähigkeit, das u. a. für Senioren besonders empfehlenswert ist.

Umgebungsbezogenes Wohlbefinden

- Eine **Landschaft** (siehe Abschnitt 7.1.1) stellt ein Gebiet dar, dessen Gelände, Vegetation und Bebauung maßgeblichen Einfluss auf Temperatur und Luftfeuchtigkeit besitzt und somit auch das umgebungsbezogene Wohlbefinden beeinflusst. Grünflächen können zudem zu einer Reduktion des Stresslevels beitragen.
- **Klanglandschaft** (siehe Abschnitt 7.1.1) kann sich sowohl positiv als auch negativ auf das menschliche Wohlbefinden auswirken. So können Naturgeräusche, wie z. B. Vogelgezwitscher, zu einem Stressabbau beitragen, eine Kreissäge oder ein Presslufthammer jedoch das Gegenteil bewirken.
- **Temperatur und Luftfeuchtigkeit** (siehe Abschnitt 7.1.1) der Umgebung besitzen einen direkten Einfluss auf die menschliche Physiologie und können daher dessen umgebungsbezogenes Wohlbefinden deutlich beeinflussen.
- Die **Luftqualität** (siehe Abschnitt 7.1.1) setzt sich aus verschiedenen Gasen sowie Feinstaub zusammen und kann negativen Einfluss auf das umgebungsbezogene Wohlbefinden ausüben. Insbesondere das Gas „Kohlendioxid“ beeinträchtigt bei hoher Konzentration in der Raumluft die kognitive Leistungsfähigkeit.

Physisches Wohlbefinden

- Die **Ernährung** (siehe Kapitel 8) besitzt einen direkten Einfluss auf den menschlichen Körper (Physis) und somit auch auf das physische Wohlbefinden. Gesunde Ernährung, bei der u. a. eine übermäßige Aufnahme von Kalorien vermieden und ausreichend getrunken wird, besitzt daher positive Effekte.
- Die **physische Fitness** (siehe Abschnitt 9.2.2) kann durch ausreichend Bewegung und körperlichen Übungen erreicht werden, die das physische Wohlbefinden verbessern. Insbesondere bei Senioren können beispielsweise bereits regelmäßige Beweglichkeitsübungen Gelenkschmerzen durch Arthrose verringern.

10.1 Wissenschaftliche Beiträge

Die in Abschnitt 1.3 vorgestellten wissenschaftlichen Forschungsziele dieser Arbeit:

- Entwicklung von Verfahren zur Erfassung von Wohlbefindensfaktoren,
- Entwicklung geeigneter Computersysteme zur Verbesserung des Wohlbefindens und
- Gewährleistung der Privatsphäre

konnten erreicht werden. Im Nachfolgenden werden die eigenen Beiträge im Hinblick auf die Forschungsziele detailliert zusammengefasst. Eine kurze Beschreibung des Anteils des Autors an den veröffentlichten Untersuchungen ist in Anhang A für jede Publikation aufgeführt. Sie beinhaltet auch den Anteil an den Hardware- / Software-Prototypen. Die Beteiligung des Autors an (mobile)SSI-Plugins ist in Anhang C detaillierter, mit Links zum Quelltext auf Github, aufgeführt.

10.1.1 Entwicklung von Verfahren zur Erfassung von Wohlbefindensfaktoren

In diesem Abschnitt werden die in dieser Arbeit vorgestellten und in wissenschaftlichen Publikationen veröffentlichten Verfahren zusammengefasst, die eine Erfassung der ausgewählten Wohlbefindensfaktoren mittels Verwendung von IoT- und Mobile-Computing-Technologie ermöglichen.

Emotionales, mentales und soziales Wohlbefinden

Die **Erkennung des Stresslevels** („entspannt“, „konzentriert“, „gestresst“) in der Büroumgebung wurde in Abschnitt 6.3.1 mithilfe einer Laborstudie untersucht. In dieser Arbeit wurden drucksensitive Matten neben etablierten physiologischen Sensoren eingesetzt. Die Studienergebnisse zeigten, dass die Daten der Matten ausreichen können, um eine Stresserkennung zu ermöglichen. Insbesondere an stationären Arbeitsplätzen kann diese Art der Sensorik daher eine geeignete, in die Umgebung integrierte Alternative oder Ergänzung zur Erkennung mit Wearables darstellen.

Wesentliche Besonderheiten dieser Untersuchung sind:

- spezielle Sensorik: Druckmatten (am Stuhl und Boden angebracht) für Stresserkennung,
- Klassifikation von drei Stressleveln (in Literatur meistens nur zwei),
- mit SSI synchronisiert im Labor aufgezeichnetes Datenkorpus (hierfür wurden vom Autor SSI-Plugins (mit)entwickelt),
- Datenanalyse mit mehreren Verfahren des maschinellen Lernens und multimodaler Fusion.

In Abschnitt 6.3.2 wurde ein Prototyp zur mobilen **Lacherkennung** („Lachen“, „nicht Lachen“) in Echtzeit mit einem Smartphone vorgestellt und evaluiert, wobei Audio- und Beschleunigungsdaten eingesetzt wurden. Es wurde hierfür ein Audiokorpus in einem Alltagsszenario aufgezeichnet. Die Klassifikationen wurden mit SVMs durchgeführt und neben unimodaler Klassifikationen zwei Fusions-Verfahren zur multimodalen Klassifikation untersucht. Insbesondere bei den multimodalen Klassifizierungen konnten vielversprechende Erkennungsraten erreicht werden. Eine Nutzung des Systems durch eine Personengruppe war mittels Vernetzung mehrerer Smartphones ebenfalls möglich. Hierdurch könnte künftig u. a. die soziale Verbundenheit untersucht werden.

Wesentliche Besonderheiten dieser Untersuchung sind:

- mobile multimodale Echtzeitsignalverarbeitung auf Smartphone mit mobileSSI,
- mittels mobileSSI im Feld aufgezeichnetes, synchronisiertes Datenkorpus (hierfür wurden vom Autor mobileSSI-Plugins (mit)entwickelt),
- Evaluation mit maschinellern Lernen mit multimodaler Fusion unter Nutzung von mobileSSI.

Umgebungsbezogenes Wohlbefinden

In der urbanen **Landschaft** existieren unterschiedliche lokale Klimazonen, die von der Bebauung und Begrünung abhängen. In Abschnitt 7.3.2 wurden die Auswirkungen von drei lokalen Klimazonen („Stadt“, „Wiese“, „Wald“) auf den Menschen in einer Feldstudie untersucht. Hierbei wurden Fragebogenantworten sowie physiologische (IBI, EDA) und Umweltdaten (u. a. Temperatur, Luftfeuchtigkeit, Audiodaten) mit mobilen Geräten aufgezeichnet und ausgewertet. Zur Erfassung der Feinstaubkonzentration, relativen Konzentration verschiedener Gase, relativen Luftfeuchtigkeit, Lufttemperatur und des Luftdrucks wurde eine vom Autor dieser Arbeit eigen entwickelte mobile Gas-Sensorbox verwendet. Die Auswirkungen der Klimazonen auf die Physiologie konnten mit den BVP- und EDA-Daten des eingesetzten Fitnessarmbands mittels Klassifizierungsmodellen nachvollzogen werden. Basierend auf der Lufttemperatur und relativen Luftfeuchtigkeit konnten die lokalen Klimazonen, insbesondere im Sommer, bereits grundsätzlich unterschieden werden. Zudem konnten mit den Audio- und physiologischen Daten Klassifizierungsmodelle zur Vorhersage der emotionalen Valenz erstellt werden.

Wesentliche Besonderheiten dieser Untersuchung sind:

- Aufzeichnung eines synchronisierten Datenkorpus verschiedener mobiler Geräte im Feld mittels mobileSSI,
- Einsatz einer speziellen, vom Autor entwickelten, mobilen Gas-Sensorbox für zusätzliche Umweltdaten,
- u. a. mittels mobileSSI durchgeführte Datenanalyse mit maschinellern Lernen und teils multimodaler Fusion,

- Beziehungen zwischen lokalen Klimazonen und Physiologie bzw. Umgebungsgeräuschen und emotionaler Valenz konnten mit der niedrigen Datenqualität der Sensorik eines Smartphones und Fitnessarmbands festgestellt werden.

Eine automatische Erfassung von **Klanglandschaften** mit tiefen neuronalen Netzen wurde in Abschnitt 7.3.1 präsentiert. Bei dieser Untersuchung lag der Fokus auf der automatischen Erstellung optimierter Netzarchitekturen mittels genetischer Algorithmen innerhalb einer akzeptablen Zeitdauer. Es wurde hierbei ein bereits existierendes Datenkorpus mit zehn verschiedenen Audioklassen, bestehend aus urbanen Geräuschen, eingesetzt. Die generierten neuronalen Netze zeigten bei der DCASE-Challenge 2018, auch im Vergleich mit von sonstigen Forschern erstellten neuronalen Netzen, gute Erkennungsraten.

Wesentliche Besonderheiten dieser Untersuchung sind:

- Entwicklung eines genetischen Algorithmus zur automatischen Erstellung von DNN-Netzarchitekturen zur Audioklassifikation,
- der Algorithmus wurde im Rahmen der DCASE-Challenge 2018 (Acoustic Scene Classification) evaluiert,
- die basierend auf den generierten Netzarchitekturen trainierten Modelle schnitten deutlich besser als die Baseline ab und übertrafen teilweise die Erkennungsraten der von sonstigen Forschern entwickelten Modelle.

Physisches Wohlbefinden: Ernährung

Zur Erfassung der Lebensmittelart und deren Menge wurden in dieser Arbeit mehrere Verfahren zum Thema Ernährung untersucht. In Abschnitt 8.3.1 wurde im Rahmen der „Paralinguistic Eating Condition (EC) Challenge“ bei der Interspeech 2015 der Einsatz von hierarchischen SVMs zur Audioklassifizierung der **Essgeräusche** von sieben verschiedenen Lebensmitteln analysiert. Durch die Identifikation und Gewichtung von für bestimmte Lebensmittelkonsistenzen markanten Audiofrequenzbereichen konnten mit dem entwickelten Verfahren insgesamt gute Erkennungsraten erreicht werden.

Wesentliche Besonderheiten dieser Untersuchung sind:

- Untersuchung von hierarchischen SVMs mit gezielter Frequenzgewichtung zur Audioklassifikation von Essgeräuschen,
- das Verfahren wurde mit SSI umgesetzt und im Rahmen der Interspeech Paralinguistic Eating Condition 2015 Challenge evaluiert.

Eine weitere Möglichkeit zur Erfassung der Lebensmittelart wurde mithilfe von Gassensorik in den Abschnitten 8.3.2 und 8.3.3 vorgestellt. Für die Untersuchungen wurden spezielle Gas-Sensorboxen mit verschiedenen Gassensoren entwickelt, die mehrere Messprinzipien (MOS, NDIR) und Fertigungstechniken (teilweise MEMS) nutzten. Diese Prototypen wurden eingesetzt, um die entstehenden **Gaskonzentrationen** von 24 Lebensmitteln (flüssig und fest) kontrolliert aufzuzeichnen. Diese Daten wurden anschließend genutzt, um Klassifizierungsmodelle zu trainieren und die Erkennungsraten verschiedener Sensorkombinationen zu untersuchen. Es zeigte sich, dass unter kontrollierten Bedingungen hohe Erkennungsraten erreicht werden und hierfür auch kompakte, stromsparende MEMS-Sensoren verwendet werden können. Ein Problem bei der Klassifizierung bestand hierbei darin, dass mit bestimmten Sensoren erstellte Modelle (sensorabhängig) nicht mit anderen Gassensoren desselben Typs eingesetzt werden konnten. Dieses Problem wurde in Abschnitt 8.3.3 behandelt und konnte mit Transferlernen weitestgehend gelöst werden.

Wesentliche Besonderheiten dieser Untersuchung sind:

- Einsatz vom Autor mitentwickelter Gas-Sensorboxen u. a. mit modernen MEMS-Gassensoren ausgestattet,

- für Untersuchung im Labor aufgezeichnetes Datenkorpus von 24 Lebensmitteln,
 - Klassifizierung von zusätzlichen Lebensmittelarten als in Literatur,
 - Analyse verschiedener Gassensorkombinationen,
 - Anwendung von Transferlernen für eine sensorunabhängige Klassifizierung.
-

Des Weiteren wurden Erfassungsverfahren untersucht, bei denen der Nutzer direkt mit dem System interagieren muss. In Abschnitt 8.3.4 wurde ein Prototyp zur Lebensmittelprotokollierung entwickelt, der aus einer Gerätekombination aus Smartphone, Smartwatch und Smartscale bestand, die mittels Bluetooth Daten austauschten. Hierdurch war eine **Multi-Geräte-Interaktion** möglich, die es Nutzern erlaubte, die jeweiligen vorteilhaften Eigenschaften (z. B. verschiedene Eingabemodalitäten) eines bestimmten Gerätes, abhängig vom Nutzungskontext, auszuschöpfen. Bei der Entwicklung wurde auf hohe Usability Wert gelegt und entsprechende Anforderungen herausgearbeitet. Es wurden für die Untersuchung zwei vom Autor entwickelte, mobile Smartscale-Prototypen eingesetzt, die eine präzise Bestimmung der Menge von Lebensmitteln erlaubten und die Messdaten an Smartphone bzw. Smartwatch übertrugen. In einer Feldstudie konnte das Multi-Geräte-Konzept überzeugen, denn es ermöglichte eine einfache Lebensmittelerfassung. Die Smartscale wurde insbesondere in privaten Situationen als sehr hilfreich angesehen, da sie eine präzise Mengenbestimmung deutlich erleichterte.

Wesentliche Besonderheiten dieser Untersuchung sind:

- mobile, multimodale Multi-Geräte-Interaktion mit vom Autor erstelltem Smartscale-Prototyp und Smartphone- und Smartwatch-App,
 - nutzerzentrierte Entwicklung,
 - Feldstudie mit Nutzern.
-

Bei einem weiteren System, der sogenannten „DrinkWatch“, die in Abschnitt 8.3.5 vorgestellt wurde, war der Nutzer in den Lernprozess der Klassifikation von **Trinkbewegungen** („trinken“, „nicht trinken“) mittels interaktivem maschinellern Lernen (iML) einbezogen. Der Prototyp bestand hierbei aus einer Smartwatch, auf der die Software zur Echtzeitsignalverarbeitung, iML und GUI ausgeführt wurde. Für die Erfassung der Armbewegungen beim Trinken wurden Accelerometer-Daten eingesetzt. Ein Basismodell wurde für eine grundsätzliche Erkennung von Trinkbewegungen (Warmstart) trainiert. Hierfür wurden in Alltagssituationen Trinkbewegungen mehrerer Personen aufgezeichnet. Die anschließende Datenannotation wurde mit einem der eigenen Smartscale-Prototypen, sowie einer Annotationssoftware mit kooperativem maschinellen Lernen erleichtert. Das aus diesen Daten erstellte Warmstartmodell wurde auf der Smartwatch mittels sporadischer Nutzerabfragen hinsichtlich dessen Trinkbewegungen optimiert. Bei diesem System konnte gezeigt werden, dass auf mobiler Hardware iML in Echtzeit ausgeführt werden kann. Bei einer Nutzerstudie konnte ein vielversprechender Anpassungsvorgang des Systems an die Trinkbewegungen der Teilnehmer beobachtet werden.

Wesentliche Besonderheiten dieser Untersuchung sind:

- mobile Echtzeitsignalverarbeitung auf Smartwatch mittels mobileSSI mit interaktivem maschinellern Lernen,
- mit mobileSSI im Feld aufgezeichnetes Datenkorpus mit Daten des Smartscale-Prototyps des Autors,
- Einsatz von kooperativem maschinellern Lernen zur Datenannotation,
- Feldstudie mit Nutzern.

10.1.2 Entwicklung geeigneter Computersysteme zur Verbesserung des Wohlbefindens

In diesem Abschnitt werden die in dieser Arbeit präsentierten und in wissenschaftlichen Publikationen veröffentlichten Assistenz-, Informations- und Empfehlungssysteme hinsichtlich ihrer Zuordnung zu Wohlbefindenskategorien zusammengefasst.

Emotionales, mentales und soziales Wohlbefinden

In Abschnitt 6.3.3 wurde ein Biofeedback-System vorgestellt, das mittels eines vom Autor entwickelten mobilen Prototyps den eigenen Herzschlag bzw. den eines Partners mit einem Tangible-Display spürbar machen konnte. Die autarken Prototypen wurden als „PiHeart“ bezeichnet und bestanden jeweils aus einem BVP-Sensor, einer Haupteinheit und einer Ausgabeeinheit. Die Haupteinheit beinhaltete u. a. einen Einplatinencomputer, der eigenentwickelte Software zur Echtzeitauswertung der Herzrate aus dem BVP-Signal sowie Ansteuerung der Ausgabeeinheit ausführte. Die Ausgabeeinheit ähnelte äußerlich einem menschlichen Herz. Herzschläge wurden elektromechanisch simuliert. Die PiHeart-Prototypen waren per WLAN ansteuerbar. Bei einer Feldstudie mit 60 Teilnehmern wurde die Wirkung des Systems, das aus einem Laptop und zwei PiHearts bestand, auf das Nutzererlebnis bei der Betrachtung ausgewählter Kurzfilme untersucht. In der Auswertung zeigte sich, dass das System das Potenzial besaß, die Selbstwahrnehmung und soziale Verbundenheit zu verbessern und somit positive Auswirkungen auf das emotionale, mentale und soziale Wohlbefinden hatte.

Wesentliche Besonderheiten dieser Untersuchung sind:

- Entwicklung eines speziellen mobilen Hardware- / Software-Prototyps (PiHeart) zur Erfassung der Herzrate und deren taktilen Ausgabe,
- Feldstudie mit 60 Teilnehmern zur Untersuchung der Auswirkungen des (geteilten) Biofeedbacks auf das Nutzererlebnis zweier Personen in einer Alltagssituation.

Umgebungsbezogenes Wohlbefinden

In Abschnitt 7.3.3 wurde ein ambientes System nutzerzentriert entwickelt, das den Bewohnern den Status von offenen- bzw. geschlossenen Fenstern und Türen mittels farbigem, ambientem Licht anzeigen konnte. Hierfür wurde ein bereits existentes Smarthome-System in einem Einfamilienhaus erweitert, indem an geeigneten Positionen wie z. B. in der Nähe der Haustür, RGBW-Leuchten installiert wurden. Die Positionierung der Leuchten sowie der Einsatz einer verständlichen Farbkodierung wurden während des Entwicklungsprozesses zusammen mit den Bewohnern festgelegt. Das ambiente Licht ermöglichte es, sofort zu erkennen, ob z. B. vor dem Verlassen des Hauses Fenster bzw. Terrassen- / Balkontüren geschlossen werden sollten. Das System wurde anschließend in einer Langzeitstudie evaluiert. Es zeigte sich hierbei, dass die Bewohner an den gewählten Leuchtmittelpositionen während der Alltagsroutinen regelmäßig vorbeikamen. Die Bewohner äußerten in einer Befragung, dass sie sich sicherer fühlten, wenn sie den Status von Fenstern und Türen einfach erkennen konnten und zudem diese wichtige Information unaufdringlich angezeigt wurde. Das System konnte somit das Sicherheitsgefühl in der Wohnumgebung verbessern und besaß daher positive Auswirkungen auf das umgebungsbezogene Wohlbefinden.

Wesentliche Besonderheiten dieser Untersuchung sind:

- Erweiterung eines bestehenden Smarthomes zu einem alltagstauglichen ambienten System zur unaufdringlichen Informationsanzeige geöffneter Fenster- / Türen,
- nutzerzentrierte Entwicklung durch Festlegung der Platzierung von RGBW-Leuchtmitteln und deren Farbkodierung zusammen mit den Bewohnern,
- Feldstudie in einem Haushalt über mehrere Monate.

Ein weiteres System, das ambientes Licht als Systemausgabe nutzte, wurde in Abschnitt 7.3.4 vorgestellt. Dieses System sollte einen Hinweis geben, wenn die Luftqualität in einem Besprechungsraum durch Lüften verbessert werden muss. Das entwickelte System nutzte hierbei einen

Einplatinencomputer, Gassensorik, Sensorik eines Smarthome-Systems sowie eigens konstruierte, per WLAN steuerbare LED-Streifen. Es wurde u. a. die eigenentwickelte, quelloffene Software „Eventerpretor“ zur ereignisbasierten Ansteuerung mit Regelsystem eingesetzt. Die Luftqualität wurde über die CO₂-Konzentration bestimmt. Die ambienten Lichter (LED-Streifen) wurden neben den Fenstergriffen platziert und die Luftqualität per Ampel-Farbkodierung signalisiert. Neben der Luftqualität zeigten die LED-Streifen den offenen bzw. geschlossenen Zustand der Fenster per Helligkeitsmuster an und erinnerten somit beim Verlassen des Raumes an das Schließen der Fenster. In einer Langzeitstudie wurden mit dem System Daten (ohne Systemausgaben) aufgezeichnet. Es konnte somit festgestellt werden, dass der Raum in einigen Fällen nicht ausreichend gelüftet wurde, sodass CO₂-Konzentrationen auftraten, die, gemäß Literatur, das Wohlbefinden der Raumnutzer reduziert haben dürften. Hiermit wurde die Notwendigkeit eines Systems zur Erinnerung an das Lüften verdeutlicht. Zusätzlich wurde ein Workshop durchgeführt, bei dem sich u. a. zeigte, dass die Lichtkodierungen von den Teilnehmern verstanden wurden und das System potenziell zur Verbesserung der Luftqualität und somit des Wohlbefindens der Raumnutzer geeignet ist.

Wesentliche Besonderheiten dieser Untersuchung sind:

- Nutzung von Smarthome-Komponenten und vom Autor entwickelten Prototypen (Sensorik / Aktuator) zur Entwicklung eines alltagstauglichen, unaufdringlichen ambienten Systems, das Nutzer beim ausreichenden Lüften von Räumen unterstützt,
- Entwicklung der quelloffenen Software „Eventerpretor“ (auf Github veröffentlicht: ¹) zur „flow-based“ Verarbeitung von ereignisbasierten Daten mit Regelsystem,
- Datenaufzeichnung / Analyse über mehrere Monate,
- Vergleich mehrerer, kostengünstiger Gassensoren (CO₂, VOCs) zur Ermittlung der Luftqualität,
- Workshop mit Raumnutzern, um Einblicke in die Verständlichkeit der ambienten Ausgaben und Ideen für Verbesserungen / Erweiterungen zu erhalten.

Physisches Wohlbefinden: Ernährung

Ein System, das mittels eines sozialen Roboters sprachliche Nährwertinformationen zu bestimmten Getränken lieferte, wurde in Abschnitt 8.3.6 vorgestellt. Der sogenannte „Drink-O-Mender“ sollte hierbei auch durch Anpassung der Sprachformulierungen, die Besucher einer Wissenschaftsausstellung überzeugen, aus fünf Getränken, eines mit möglichst geringem Zuckergehalt auszuwählen, das sie anschließend in ein Trinkgefäß gießen und trinken konnten. Das System basierte auf der vom Autor entwickelten Hardware, die zunächst erkannte, ob ein Nutzer präsent war, welches Getränk er auswählte und wie viel er trank. Für die Mengenerfassung wurde hierbei eine angepasste Smartscale eingesetzt, die in diesem Fall über LED-Beleuchtung verfügte und den Ablauf verdeutlichte. Für die eigene Hardware wurde vom Autor entsprechende Software entwickelt und Eventerpretor für die Datenauswertung und Logik eingesetzt. Beobachtungen während der Systemnutzung bei einer Ausstellung zeigten, dass das System bei den meisten Besuchern positiv angenommen wurde, der Roboter die Besucher jedoch nicht immer von einem gesünderen Getränk überzeugen konnte. Das System war sehr wohl in der Lage, Personen auf ungesunde Getränke aufmerksam zu machen und damit zu einer bewussten Entscheidungsfindung bei der Getränkewahl beizutragen.

Wesentliche Besonderheiten dieser Untersuchung sind:

- Prototyp (Drink-O-Mender) bestehend aus einem sozialen Roboter sowie vom Autor entwickelter Getränkekartonsensoren und angepasster Smartscale zur Unterstützung bei einer gesundheitsorientierten Getränkewahl,
- Verarbeitung der Sensordaten mittels Eventerpretor,
- Anpassung der Sprachformulierungen des Roboters an die Nutzer mittels Reinforcement Learning,

¹<https://github.com/andreas-seiderer/Eventerpretor>

- alltagstauglicher, robuster Prototyp, der in Feldstudie von 78 Besuchern einer Ausstellung verwendet wurde.

Zum Thema Ernährung wurde in Abschnitt 8.3.7 ein weiteres Informationssystem vorgestellt. Es handelte sich hierbei um einen Sprachassistenten, der es ermöglichte, von verpackten Lebensmitteln per Barcode und lokaler Datenbank Nährwertinformationen abzufragen. Zudem konnte ausgegeben werden, ob ein Lebensmittel gesund ist (Nutri-Score) oder z. B. der Zuckergehalt zu hoch ist. Der Sprachassistent konnte per Webinterface verwendet werden, das neben natürlicher Sprachein- und -ausgaben, textuelle Interaktion erlaubte. Das System war vollständig ohne Internetverbindung nutzbar und alle Systemkomponenten waren quelloffen. Hierzu zählten auch die deutsche Spracherkennung, Sprachverarbeitung und Sprachausgabe. Die Arbeit gibt einen aktuellen Überblick über für Sprachassistenten geeignete, quelloffene Komponenten. In einem Workshop mit zwei Senioren wurde die Verwendbarkeit des Sprachassistenten mit Nutzung eines Tablet untersucht. Die wesentlichen Ergebnisse waren, dass ungesunde Lebensmittel relativ einfach identifiziert werden konnten und für Senioren, ohne Lesehilfe / Brille, viele klein gedruckte Informationen auf Lebensmittelverpackungen leicht zugänglich waren. Somit konnte der Sprachassistent einen Beitrag für eine gesündere Ernährung leisten.

Wesentliche Besonderheiten dieser Untersuchung sind:

- Entwicklung eines deutschen, offlinefähigen Sprachassistenten zur Ausgabe von (interpretierten) Nährwertinformationen verpackter Lebensmittel, der aus quelloffenen Softwarekomponenten bestand,
- Überblick über für Sprachassistenten geeignete, quelloffene Komponenten,
- Workshop mit zwei Senioren, um die praktische Verwendbarkeit des Sprachassistenten für diese Zielgruppe zu untersuchen.

Emotionales, mentales, soziales, umgebungsbezogenes, physisches Wohlbefinden

In Kapitel 9 wurden das CARE-System und hierauf basierende Weiterentwicklungen vorgestellt. Beim CARE-System handelte es sich um ein kontextsensitives Empfehlungssystem für Senioren, das einen maßgeblichen Beitrag zur Verbesserung ihres Wohlbefindens erbringen sollte. Das System beinhaltete Empfehlungen zu allen Kategorien des Wohlbefindensmodells. Die Grundidee bestand darin, ein System zu entwickeln, das unaufdringlich in die Wohnung integrierbar ist, und Senioren zum aktuellen Kontext passende Empfehlungen erhalten können. Die einzelnen Phasen des nutzerzentrierten Entwicklungsprozesses wurden in der Arbeit ausführlich beschrieben. Hierbei entstanden drei Prototypen, die vom Autor entwickelte Hard- und Software sowie kommerzielle Smarthome-Hardware einsetzten. Im Rahmen des nutzerzentrierten Entwicklungsprozesses wurden bei jeder Systemiteration Verbesserungen und Anpassungen, basierend auf vorherigen Nutzerevaluationen, berücksichtigt. Hierdurch konnte ein System entwickelt werden, das von Senioren einfach nutzbar war und regelmäßig verwendet wurde. Um das Interesse einer Langzeitnutzung zu erhöhen, wurden in den finalen CARE-Prototyp Gamifizierungen integriert. Mittels zweier Workshops mit Senioren eines Altenheims und Seniorentreffs wurde das Belohnungsschema der Gamifizierungen erarbeitet. In einer Studie mit dem weiterentwickelten CARE-Prototyp führten u. a. die Gamifizierungen zu einem aktiveren Nutzerverhalten.

Um das Interesse am System nachhaltig zu steigern, wurde im Rahmen eines ergänzenden Projekts ein sozialer Roboter integriert. Er konnte Informationen per Sprache ausgeben und Formulierungen an den jeweiligen Nutzer anpassen. Dies erfolgte durch Reinforcement Learning, basierend auf dem Nutzerfeedback mittels eines eigenentwickelten Bedienpults. Eine Evaluation mit Senioren zeigte, dass sie, je nach Anwendungsszenario, die unterschiedlichen Sprachformulierungen des Roboters bevorzugten. Infolge der Verwendung des Bedienpults besaß das System eine hohe Usability. Die Senioren bewerteten das System insgesamt positiv und attraktiv.

Wesentliche Besonderheiten des CARE-Projekts sind:

- Entwicklung eines alltagstauglichen, in die Wohnung integrierten Prototyps zur Verbesserung des Wohlbefindens von Senioren mittels kontextsensitiver visueller Empfehlungen,
- automatischer Wechsel zwischen ambientem Bilderrahmen-Modus (Nutzer in Systemnähe) in interaktiven Empfehlungs-Modus (Nutzer steht direkt vor System),
- Entwicklung eines pragmatischen Wohlbefindensmodells bestehend aus fünf Kategorien,
- insgesamt wurden ca. 130 Empfehlungen (CARE-Prototyp 2) zu entsprechenden Wohlbefindenskategorien erstellt und abhängig vom Kontext ausgegeben,
- Erfassung mehrerer Wohlbefindensfaktoren und Kontextelemente mittels unaufdringlicher Sensoren und Nutzereingaben,
- nutzerzentrierte Entwicklung über drei Iterationen mit Senioren,
- Berücksichtigung der Anforderungen von Senioren, z. B. Hardwaretaster als Alternative zur Eingabe per Touchscreen,
- Untersuchung und Umsetzung von Gamifizierungen u. a. mit per Fitnessarmband und drucksensitiver Bodenmatte überwachten Übungen zur Steigerung der Langzeitmotivation,
- Feldstudien in der Wohnung von zwei Senioren mit den Systemprototypen über mehrere Wochen.

Wesentliche Besonderheiten des weiterführenden Projekts mit einem sozialen Roboter sind:

- Nutzung eines sozialen Roboters, der bei verschiedenen Anwendungen seine Sprachformulierungen an den Nutzer anpassen konnte,
- Nutzung der Empfehlungen des CARE-Projekts,
- Einsatz eines vom Autor dieser Arbeit entwickelten Bedienpanels für eine erleichterte Systemeingabe,
- Feldstudien in der Wohnung von Senioren.

10.1.3 Gewährleistung der Privatsphäre

Um den Schutz der Privatsphäre zu gewährleisten, wurde dies von Beginn an bei der Entwicklung der Systeme berücksichtigt. Das war deshalb erforderlich, da bei Systemen zur Verbesserung des menschlichen Wohlbefindens persönliche und vor allem auch gesundheitsbezogene Daten verarbeitet werden müssen. Entsprechend wurden in dieser Arbeit ausgewählte Sensoren eingesetzt und zudem die Echtzeitsignalverarbeitung sowie das maschinelle Lernen auf Nutzergeräten u. a. mit (mobile)SSI und Eventerpretor ausgeführt, womit vermieden werden konnte, dass schützenswerte Daten per Internet an externe Anbieter übertragen und von diesen gespeichert bzw. verarbeitet werden. Eine Umsetzung ist nicht einfach möglich, da z. B. moderne, tiefe neuronale Netze in vielen Fällen nicht trainiert werden können, wenn keine ausreichend hohe Datenmenge oder Rechenleistung vorhanden sind. Eine Lösung kann Transferlernen darstellen, das in der vorliegenden Arbeit in Abschnitt 5.5.2 untersucht wurde. Es konnte gezeigt werden, dass mit dieser Technik auf Smartphones und Embedded Systemen tiefe neuronale Netze mit relativ hohen Erkennungsraten in akzeptabler Zeit und mit kleinen Datenmengen trainiert werden können. Dies könnte beispielsweise für eine personalisierte Smartphone-App zur Bilderkennung eingesetzt werden.

Wesentliche Besonderheiten der Untersuchung aus Abschnitt 5.5.2 sind:

- Untersuchung der praktischen Verwendbarkeit von Transferlernen zum Trainieren von tiefen neuronalen Netzen zur Bildklassifizierung auf mobilen / Embedded Geräten,
- Analyse der Trainingszeiten auf mehreren Geräten und Untersuchung der Erkennungsraten bei niedriger (unausgeglichener) Bildanzahl pro Klasse,
- Veröffentlichung der verwendeten Programme und Ergebnisse auf Github: ².

²<https://github.com/hcmlab/tf-transferlearning-benchmarks> (08.05.2019)

10.2 Fortführende Arbeiten

Viele Möglichkeiten für eine Vertiefung und Weiterführung von ausgewählten Verfahren, Untersuchungen und Systementwicklungen wurden bereits in dieser Arbeit angesprochen. Beispielsweise sind beim umfassenden CARE-System eine Vielzahl an Erweiterungen und Untersuchungen bezüglich der Sensorik zur Kontexterfassung sowie der Auswahl und Ausgabe von Empfehlungen möglich. Insbesondere bei der Erfassung von Wohlbefindensfaktoren wurden in dieser Arbeit bereits diverse Möglichkeiten aufgezeigt, die zu verschiedenen Forschungsbereichen, wie z. B. der Erkennung von menschlichen Aktivitäten (human activity recognition), zählen und weiterhin sehr aktiv untersucht werden. Es ist zudem zu beobachten, dass die Technikaffinität von älteren Senioren kontinuierlich zunimmt und diese immer häufiger das Internet nutzen (siehe hierzu den D21-Digital-Index). Von einer alle Altersgruppen umfassenden höheren Akzeptanz für entsprechende Computersysteme ist daher in Zukunft auszugehen. Insbesondere, die derzeit besonders populären kommerziellen Sprachassistenten von Amazon, Apple oder Google, die teilweise bereits mit einem Smarthome-System, Touchscreen und mobilen Geräten kombiniert sind, könnten für umfassendere Systeme zur Verbesserung des Wohlbefindens eingesetzt werden. In dieser Arbeit konnte bereits gezeigt werden, dass hierfür grundsätzlich auch quelloffene Lösungen genutzt werden können, bei denen der Datenschutz garantiert werden kann. Eine Kombination mit sich an die Nutzer anpassenden, sozialen Robotern wäre ebenfalls vorstellbar und könnte sich beispielsweise positiv auf die Attraktivität und Langzeitnutzung von entsprechenden Systemen auswirken. Beim Design der Systeminteraktion sollte jedoch berücksichtigt werden, dass sich diese, je nach Anwendung und Situation, im Aufmerksamkeitshintergrund der Nutzer abspielen sollte. Ansonsten könnten diese schnell störend wirken oder Nutzer überfordern.

Im Bereich des maschinellen Lernens für Verfahren und Systeme zur Verbesserung des Wohlbefindens sind diverse fortführende Arbeiten bereits absehbar. Tiefe neuronale Netze werden Dank neuer Hardware und optimierter Software im IoT- und Mobile-Computing-Bereich immer häufiger angewandt und zeigen hierbei sehr gute Leistung. Es werden zudem zunehmend leistungsfähigere Verfahren, wie „Generative Adversarial Networks“ (GAN) (Goodfellow et al., 2014) entwickelt und neue Einsatzmöglichkeiten erforscht. Schlüsseltechnologien wie Transferlernen (siehe Abschnitt 5.5.2) ermöglichen die Nutzung tiefer neuronaler Netze auch mit nur geringen Datenmengen. Da tiefe neuronale Netze jedoch sehr komplex und hierdurch für den Menschen nicht bzw. nur schwer verständlich sind, wird derzeit an Techniken wie Explainable AI (XAI) (siehe Literaturüberblick von Adadi und Berrada (2018)) geforscht, die u. a. dabei helfen können, fehlerhafte Modelle zu erkennen. Dies kann auch für die automatische Erstellung von tiefen neuronalen Netzen hilfreich sein, wodurch auch, ohne manuelles Eingreifen durch Experten, passende Netzarchitekturen sowie Hyperparameter bestimmt werden können. Ohne XAI existieren bereits kommerzielle Produkte wie Googles AutoML³ zur automatischen Erzeugung von Modellen. Ein Problem hierbei bleibt allerdings, dass beschriftete Daten benötigt werden. Dies ist teilweise auch automatisch möglich, indem Verfahren des unüberwachten maschinellen Lernens oder (vorübergehend) andere Sensoren eingesetzt werden. In einigen Fällen wird jedoch der Nutzer für Datenbeschriftungen benötigt. In diesem Fall kann interaktives und kooperatives maschinelles Lernen eingesetzt werden, wenn hierbei der Nutzer nicht zu sehr gestört oder der Aufwand zu hoch wird. Ansätze für diese Verfahren sind bereits in dieser Arbeit enthalten.

Insgesamt zeichnet sich ab, dass künftige wissenschaftliche Forschung im Bereich der IoT- und Mobile-Computing-Technologien zu ubiquitären Systemen führen kann, die einen maßgebenden Beitrag zu einer weiteren Verbesserung des menschlichen Wohlbefindens leisten könnten.

³<https://cloud.google.com/automl> (21.07.2020)

A Publikationen und Reviewaktivitäten


A.1 Wissenschaftliche Veröffentlichungen

Publikationen (peer-reviewed) mit Eigenbeteiligung sortiert nach Kapiteln und auftreten:

Kapitel 5: Transferlernen

- Seiderer, A., Dietz, M., Aslan, I. & André, E. (2018). Enabling Privacy with Transfer Learning for Image Classification DNNs on Mobile Devices. *Proceedings of the 4th EAI International Conference on Smart Objects and Technologies for Social Good*, 25–30. <https://doi.org/10.1145/3284869.3284893>
- Schiller, D., Huber, T., Lingenfelder, F., Dietz, M., Seiderer, A. & André, E. (2019). Relevance-based feature masking: Improving neural network based whale classification through explainable artificial intelligence. *20th Annual Conference of the International Speech Communication Association INTERSPEECH*

Kapitel 6: Emotionales, mentales und soziales Wohlbefinden

- Nakashima, Y., Kim, J., Flutura, S., Seiderer, A. & André, E. (2016). Stress Recognition in Daily Work. In S. Serino, A. Matic, D. Giakoumis, G. Lopez & P. Cipresso (Hrsg.), *Pervasive Computing Paradigms for Mental Health* (S. 23–33). Springer International Publishing. https://doi.org/10.1007/978-3-319-32270-4_3
- Flutura, S., Wagner, J., Lingenfelder, F., Seiderer, A. & André, E. (2016a). Laughter detection in the wild: demonstrating a tool for mobile social signal processing and visualization. *Proceedings of the 18th ACM International Conference on Multimodal Interaction - ICMI 2016*, 406–407. <https://doi.org/10.1145/2993148.2998524>
- Flutura, S., Wagner, J., Lingenfelder, F., Seiderer, A. & André, E. (2016c). MobileSSI: asynchronous fusion for social signal interpretation in the wild. *Proceedings of the 18th ACM International Conference on Multimodal Interaction - ICMI 2016*, 266–273. <https://doi.org/10.1145/2993148.2993164>
- Flutura, S., Wagner, J., Lingenfelder, F., Seiderer, A. & André, E. (2016b). MobileSSI - A Multi-modal Framework for Social Signal Interpretation on Mobile Devices. *2016 12th International Conference on Intelligent Environments (IE)*, 210–213. <https://doi.org/10.1109/IE.2016.47>
- Aslan, I., Seiderer, A., Dang, C. T., Rädler, S. & André, E. (2020b). Resonating experiences of self and others enabled by a tangible somaesthetic design. *arXiv preprint arXiv:2005.02304*
 ausführlichere Veröffentlichung der nachfolgenden Publikation auf arXiv.org
- Aslan, I., Seiderer, A., Dang, C. T., Rädler, S. & André, E. (2020a). PiHearts: Resonating Experiences of Self and Others Enabled by a Tangible Somaesthetic Design. *Proceedings of the 2020 International Conference on Multimodal Interaction*, 433–441. <https://doi.org/10.1145/3382507.3418848>

Kapitel 7: Umgebungsbezogenes Wohlbefinden

- Roletscheck, C., Watzka, T., Seiderer, A., Schiller, D. & André, E. Using an evolutionary approach to explore convolutional neural networks for acoustic scene classification. In: *Proceedings of the Detection and Classification of Acoustic Scenes and Events 2018 Workshop (DCASE2018)*. 2018, 158–162
- Flutura, S., Seiderer, A., Aslan, I., Dietz, M., Schiller, D., Beck, C., Rathmann, J. & André, E. (2019). Mobile Sensing for Wellbeing Estimation of Urban Green Using Physiological Signals. *Proceedings of the 5th EAI International Conference on Smart Objects and Technologies for Social Good*, 249–254. <https://doi.org/10.1145/3342428.3342655>
- Rathmann, J., Beck, C., Flutura, S., Seiderer, A., Aslan, I. & André, E. (2020). Towards quantifying forest recreation: Exploring outdoor thermal physiology and human well-being along exemplary pathways in a central European urban forest (Augsburg, SE-Germany). *Urban Forestry & Urban Greening*, 49, 126622. <https://doi.org/https://doi.org/10.1016/j.ufug.2020.126622>

- **Seiderer, A.**, Dang, C. T. & André, E. (2017). Exploring Opportunistic Ambient Notifications in the Smart Home to Enhance Quality of Live. In M. Mokhtari, B. Abdulrazak & H. Aloulou (Hrsg.), *Enhanced Quality of Life and Smart Living* (S. 151–160). Springer International Publishing. https://doi.org/10.1007/978-3-319-66188-9_13
- **Seiderer, A.**, Aslan, I., Dang, C. T. & André, E. (2019). Indoor Air Quality and Wellbeing - Enabling Awareness and Sensitivity with Ambient IoT Displays. In I. Chatzigiannakis, B. De Ruyter & I. Mavrommati (Hrsg.), *Ambient Intelligence* (S. 266–282). Springer International Publishing

Kapitel 8: Physisches Wohlbefinden: Ernährung

- Wagner, J., **Seiderer, A.**, Lingenfelder, F. & André, E. (2015). Combining Hierarchical Classification with Frequency Weighting for the Recognition of Eating Conditions. *INTERSPEECH*, 5
- Dang, C. T., **Seiderer, A.** & André, E. (2018). Theodor: A Step Towards Smart Home Applications with Electronic Noses. *Proceedings of the 5th international Workshop on Sensor-based Activity Recognition and Interaction*, 11. <https://doi.org/10.1145/3266157.3266215>
- **Seiderer, A.**, Dang, C. T. & André, E. (2019). Calibration Transfer for Food Recognition Models for E-Noses. *Proceedings of the 9th International Conference on the Internet of Things*, 36:1–36:4. <https://doi.org/10.1145/3365871.3365910>
- **Seiderer, A.** & André, E. (2017). Development of a Multi-Device Nutrition Logging Prototype Including a Smartscale. *Proceedings of the 2017 International Conference on Digital Health - DH '17*, 239–240. <https://doi.org/10.1145/3079452.3079486>
- **Seiderer, A.**, Flutura, S. & André, E. (2017). Development of a mobile multi-device nutrition logger. *Proceedings of the 2nd ACM SIGCHI International Workshop on Multisensory Approaches to Human-Food Interaction*, 5–12. <https://doi.org/10.1145/3141788.3141790>
- Flutura, S., **Seiderer, A.**, Aslan, I., Dang, C. T., Schwarz, R., Schiller, D. & André, E. (2018). DrinkWatch: A Mobile Wellbeing Application Based on Interactive and Cooperative Machine Learning. *Proceedings of the 2018 International Conference on Digital Health*, 65–74. <https://doi.org/10.1145/3194658.3194666>
🏆 best student paper
- Ritschel, H., **Seiderer, A.**, Janowski, K., Aslan, I. & André, E. (2018). Drink-O-Mender: An Adaptive Robotic Drink Adviser. *Proceedings of the 3rd International Workshop on Multisensory Approaches to Human-Food Interaction*, 3. <https://doi.org/10.1145/3279954.3279957>
- **Seiderer, A.**, Ritschel, H. & André, E. (2020). Development of a Privacy-By-Design Speech Assistant Providing Nutrient Information for German Seniors. *Proceedings of the 6th EAI International Conference on Smart Objects and Technologies for Social Good*, 114–119. <https://doi.org/10.1145/3411170.3411227>

Kapitel 9: Projekt CARE: Verbesserung des Wohlbefindens von Senioren

- Hammer, S., **Seiderer, A.**, André, E., Rist, T., Kastrinaki, S., Hondrou, C., Raouzaïou, A., Karpouzis, K. & Kollias, S. (2015). Design of a lifestyle recommender system for the elderly: requirement gatherings in Germany and Greece. *Proceedings of the 8th ACM International Conference on Pervasive Technologies Related to Assistive Environments - PETRA '15*, 1–8. <https://doi.org/10.1145/2769493.2769559>
- Rist, T., **Seiderer, A.**, Hammer, S., Mayr, M. & André, E. (2015). CARE - extending a digital picture frame with a recommender mode to enhance well-being of elderly people. *2015 9th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth)*, 112–120. <https://doi.org/10.4108/icst.pervasivehealth.2015.259255>
- **Seiderer, A.**, Hammer, S., André, E., Mayr, M. & Rist, T. (2015). Exploring Digital Image Frames for Lifestyle Intervention to Improve Well-being of Older Adults. *Proceedings of the 5th International Conference on Digital Health 2015 - DH '15*, 71–78. <https://doi.org/10.1145/2750511.2750514>
- Rist, T., **Seiderer, A.** & André, E. (2018). Providing Life-Style-Intervention to Improve Well-Being of Elderly People. *International Conference on Entertainment Computing*, 362–367. https://doi.org/10.1007/978-3-319-99426-0_45
- Herpich, M., Rist, T., **Seiderer, A.** & André, E. (2017). Towards a Gamified Recommender System for the Elderly. *Proceedings of the 2017 International Conference on Digital Health - DH '17*, 211–215. <https://doi.org/10.1145/3079452.3079500>

- Ritschel, H., **Seiderer, A.**, Janowski, K., Wagner, S. & André, E. (2019). Adaptive Linguistic Style for an Assistive Robotic Health Companion Based on Explicit Human Feedback. *Proceedings of the 12th ACM International Conference on Pervasive Technologies Related to Assistive Environments*, 247–255. <https://doi.org/10.1145/3316782.3316791>
🏆 best technical paper
- Ritschel, H., Janowski, K., **Seiderer, A.**, Wagner, S. & André, E. (2019). Insights on Usability and User Feedback for an Assistive Robotic Health Companion with Adaptive Linguistic Style. *Proceedings of the 12th ACM International Conference on Pervasive Technologies Related to Assistive Environments*, 319–320. <https://doi.org/10.1145/3316782.3322737>
🏆 best poster

Publikationen, die in dieser Arbeit nicht explizit behandelt werden

- Dang, C. T., Aslan, I., **Seiderer, A.** & André, E. (2020). Exploring Ambient Air Quality Notifications for Smart Rooms, 102–107. <https://doi.org/10.1145/3411170.3411229>
- Faltaous, S., Haas, G., Barrios, L., **Seiderer, A.**, Rauh, S. F., Chae, H. J., Schneegass, S. & Alt, F. (2019). BrainShare: A Glimpse of Social Interaction for Locked-in Syndrome Patients. *CHI'19 Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3290607.3312754>
🏆 basierend auf dem „BR41N.IO Brain-Computer Interface Designers Hackathon 2018“ bei der „UIST School“, bei dem der dritte Platz vom Team erreicht wurde
- Flutura, S., **Seiderer, A.**, Huber, T., Weitz, K., Aslan, I., Schlagowski, R., André, E. & Rathmann, J. (2020). Interactive Machine Learning and Explainability in Mobile Classification of Forest-Aesthetics. *Proceedings of the 6th EAI International Conference on Smart Objects and Technologies for Social Good*, 90–95. <https://doi.org/10.1145/3411170.3411225>
- Ritschel, H., Janowski, K., **Seiderer, A.** & André, E. (2019). Towards a Robotic Dietitian with Adaptive Linguistic Style. *Proceeding of the Poster and Workshop Sessions of Aml-2019, the 2019 European Conference on Ambient Intelligence*
- Ritschel, H., Aslan, I., Mertes, S., **Seiderer, A.** & André, E. (2019). Personalized synthesis of intentional and emotional non-verbal sounds for social robots. *8th International Conference on Affective Computing & Intelligent Interaction (ACII 2019)*
- Ritschel, H., **Seiderer, A.** & André, E. (2020). Pianobot: An Adaptive Robotic Piano Tutor. *Workshop on Exploring Creative Content in Social Robotics at HRI 2020*
- Wagner, J., Fraga-Silva, T., Josse, Y., Schiller, D., **Seiderer, A.** & André, E. (2017). Infected Phonemes: How a Cold Impairs Speech on a Phonetic Level. *INTERSPEECH*, 3457–3461
- Wagner, J., Schiller, D., **Seiderer, A.** & André, E. (2018). Deep Learning in Paralinguistic Recognition Tasks: Are Hand-crafted Features Still Relevant? *Proc. Interspeech 2018*, 147–151

Wissenschaftliche Publikationen an denen der Autor dieser Arbeit federführend oder mit Beiträgen beteiligt war, werden nachfolgend aufgelistet und der eigene Beitrag kurz erläutert. Die vorherige Übersicht enthält die ausführlichen Literaturangaben der eigenen Arbeiten. Thematisch überlappende Arbeiten werden gesammelt behandelt.

► **Hammer, Seiderer, André, Rist, Kastrinaki, Hondrou, Raouzaïou, Karpouzis und Kollias (2015), Rist, Seiderer, Hammer, Mayr und André (2015) und Seiderer, Hammer, André, Mayr und Rist (2015)**

Diese Arbeiten wurden im Rahmen des CARE-Projekts (BMBF) publiziert. Der Autor war an folgenden Teilen der Projektumsetzung maßgeblich beteiligt: Konzepterarbeitung, Projektplanung, Implementierung und Integration von Software (die Erfassung / Verarbeitung von Sensordaten und Interpretation von Kontextinformationen wurden ausschließlich vom Autor entwickelt), Auswahl von Hardwarekomponenten und Entwicklung des Hardware-Prototyps (die elektronischen Schaltungen wurden ausschließlich vom Autor entwickelt), Erstellung von Inhalten (z. B. geeignete Empfehlungen), Durchführung von Systemtests, Vorbereitung / Durchführung / Auswertung der Feldstudien und Erstellung der Publikationen. (**Seiderer et al.**, 2015) und ein Teil der Publikation

(Hammer et al., 2015) wurden vom Autor auf den Konferenzen präsentiert. Im Rahmen des Projekts wurden vom Autor folgende SSI-Plugins (siehe Anhang C) (mit)entwickelt: *mqtt*, *xmpp*, *websocket*, *wax9*, *generic_serial*, *winsensors*, *FileBinaryLZ4*.

► **Wagner, Seiderer, Lingenfelser und André (2015)**

Die maßgebenden Beiträge des Autors an der Arbeit sind: Unterstützung bei der Entwicklung des Konzepts, der Durchführung der Experimente sowie der Publikation.

► **Flutura, Wagner, Lingenfelser, Seiderer und André (2016a, 2016b, 2016c)**

In diesen Arbeiten wurde das Framework mobileSSI (siehe Abschnitt 4.1.2), das auf SSI (siehe Abschnitt 4.1.1) basiert, mit dem Anwendungsfall „mobile Lacherkennung“, präsentiert. Der Autor war bei diesen Arbeiten an der Softwareentwicklung (einige der verwendeten (mobile)SSI-Plugins; siehe Anhang C: *androidSensors*, *websocket*) und den Publikationen beteiligt.

► **Nakashima, Kim, Flutura, Seiderer und André (2016)**

Der Autor war an folgenden Teilen der Arbeit wesentlich beteiligt: Literaturrecherche, Sensorauswahl, Sensorbeschaffung, Sensor- und Feature-Integration in das SSI-Framework (siehe Anhang C: *sensingtex*, *touchmouse*, *pressuremat*), Vorbereitung / Durchführung der Studie und der Publikation.

► **Herpich, Rist, Seiderer und André (2017)**

Maßgebende Beiträge des Autors an dieser Arbeit bestanden darin den CARE-Prototyp 2 für die Workshops zu erweitern und vorzubereiten. Hierzu zählten die Entwicklung von Übungen (Druckmatte / Fitnessarmband). Hierfür wurden vom Autor folgende SSI-Plugins (siehe Anhang C) entwickelt: *rawinput*, *eventintpreter*. Zudem war der Autor an der Durchführung der beiden Workshops und der Publikation beteiligt. Das Papier wurde vom Autor auf der Konferenz präsentiert.

► **Seiderer und André (2017) und Seiderer, Flutura und André (2017)**

Der Autor war an diesen Arbeiten an allen Teilen maßgebend beteiligt. Die mobilen Smartscale-Prototypen 2 und 3 (mit Wägezelle und Bluetooth) wurden ausschließlich, die Smartphone- und Smartwatch-Apps primär vom Autor entwickelt. Der Autor war federführend an den Publikationen beteiligt. Die Arbeiten wurden vom Autor auf den Konferenzen präsentiert.

► **Seiderer, Dang und André (2017)**

Der Autor war an allen Teilen der Arbeit maßgebend beteiligt. Das System wurde ausschließlich vom Autor entwickelt. Der Autor war federführend an den Publikationen beteiligt. Die Arbeit wurde vom Autor auf der Konferenz präsentiert.

► **Wagner, Fraga-Silva, Josse, Schiller, Seiderer und André (2017)**

Der Autor beteiligte sich an der Arbeit bei der Durchführung zusätzlicher Experimente.

► **Rist, Seiderer und André (2018)**

Diese Arbeit ist eine Zusammenfassung des gesamten CARE-Projekts und beinhaltet die Erweiterung durch Gamifizierungen. Sie basiert grundsätzlich auf den Arbeiten (Hammer et al., 2015; Herpich et al., 2017; Rist et al., 2015; Seiderer et al., 2015), an denen der Autor, wie bereits erwähnt, maßgebend beteiligt war.

► Roletscheck, Watzka, Seiderer, Schiller und André (2018)

Die in dieser Arbeit eingesetzten Verfahren basieren auf den Ergebnissen der Masterarbeiten von Herrn Roletscheck und Watzka, die beide vom Autor betreut wurden (siehe Anhang B). Der Autor beteiligte sich unterstützend bei der Konzepterarbeitung, der Durchführung der Experimente sowie der Publikation. Die Publikation wurde vom Autor beim Workshop präsentiert.

► Flutura, Seiderer, Aslan, Dang, Schwarz, Schiller und André (2018)

Der Autor war an folgenden Teilen der Arbeit wesentlich beteiligt: Auswahl geeigneter Smartwatch-Hardware, Integration des Smartscale-Prototyp 2 / 3 des Autors (Seiderer & André, 2017; Seiderer, Flutura & André, 2017) zur Datenaufzeichnung mit mobileSSI (Verwendung des vom Autor mitentwickelten Plugins *androidSensors*), Unterstützung bei der Aufzeichnung des Datenkorpus und der Publikation.

► Dang, Seiderer und André (2018)

Der Autor war an der Prototyp-Entwicklung, dem Konzept und Durchführung der Experimente, deren Auswertung sowie der Publikation maßgeblich beteiligt.

► Ritschel, Seiderer, Janowski, Aslan und André (2018)

Der Autor hat die Getränkekartonsensoren und die Smartscale sowie die Software zu deren Auswertung (Eventpreteror) entwickelt. Bei der Präsentation des Prototyps bei einer Wissenschaftsausstellung (bei der das Experiment durchgeführt wurde) und an der Publikation war er beteiligt.

► Seiderer, Dietz, Aslan und André (2018)

Der Autor war an allen Teilen der Arbeit maßgeblich beteiligt. Die Experimente wurden vom Autor durchgeführt und ausgewertet. Der Autor war federführend an der Publikation beteiligt. Die Arbeit wurde vom Autor auf der Konferenz präsentiert.

► Wagner, Schiller, Seiderer und André (2018)

Der Autor beteiligte sich an der Arbeit im Rahmen der Durchführung zusätzlicher Experimente.

► Faltaous, Haas, Barrios, Seiderer, Rauh, Chae, Schneegass und Alt (2019)

Das Papier entstand als Resultat des „BR41N.IO Brain-Computer Interface Designers Hackathon 2018“ bei der „UIST School“. Hierbei wurde vom Team der dritte Platz erreicht. Der Autor war am Konzept und der Umsetzung des Prototyps sowie der Publikation beteiligt.

► Ritschel, Janowski, Seiderer und André (2019)

Der Autor war an Teilen der Publikation beteiligt und präsentierte die Arbeit auf der Konferenz.

► Ritschel, Aslan, Mertes, Seiderer und André (2019)

Der Autor war an der Hardware des Prototyps beteiligt.

► Ritschel, Janowski, Seiderer, Wagner und André (2019) und Ritschel, Seiderer, Janowski, Wagner und André (2019)

An dieser Arbeit wirkte der Autor durch folgende Beiträge maßgebend mit: Bereitstellung der Empfehlungsinhalte des CARE-Systems, Entwicklung der Hard- und Software des Bedienpanels (ausschließlich durch den Autor), Unterstützung bei der Integration des Bedienpanels in das System und bei der Publikation.

► Schiller, Huber, Lingenfelser, Dietz, Seiderer und André (2019)

Der Autor war an folgenden Teilen der Arbeit beteiligt: Durchführung eigener Experimente und Unterstützung bei der Umsetzung des publizierten Verfahrens.

► Seiderer, Aslan, Dang und André (2019)

Der Autor war an allen Teilen der Arbeit maßgeblich beteiligt. Der Prototyp (Hardware / Software) wurde primär vom Autor entwickelt, wobei die ausschließlich vom Autor entwickelte Software „Eventerpretor“ eingesetzt wurde. Der Autor war federführend bei der Publikation. Die Arbeit wurde vom Autor auf der Konferenz präsentiert.

► Seiderer, Dang und André (2019)

Der Autor war an allen Teilen der Arbeit maßgeblich beteiligt. Der Autor war bei der Publikation federführend.

► Flutura, Seiderer, Aslan, Dietz, Schiller, Beck, Rathmann und André (2019) und Rathmann, Beck, Flutura, Seiderer, Aslan und André (2020)

Der Autor war an folgenden Teilen der Arbeit maßgeblich beteiligt: Konzepterarbeitung, Planung und Umsetzung eines mobilen Prototyps zur Messung von Umwelteinflüssen (die mobile Gas-Sensorbox wurde ausschließlich vom Autor entwickelt), Software des Hardware-Prototyps und Anbindung an mobileSSI, Vorbereitung / Durchführung / Auswertung der Feldstudien (u. a. Berechnung der Audiofeatures) und Verfassen der Publikationen. Es wurden vom Autor (mit)entwickelte mobileSSI-Plugins (*androidSensors*, *websocket*) verwendet.

► Flutura, Seiderer, Huber, Weitz, Aslan, Schlagowski, André und Rathmann (2020)

Der mobile Software-Prototyp, der Transferlernen einsetzte, wurde primär vom Autor entwickelt. Der Autor war maßgeblich an der Publikation beteiligt.

► Seiderer, Ritschel und André (2020)

Der Autor war an allen Teilen der Arbeit maßgeblich beteiligt. Teile des Prototyps wurden ausschließlich vom Autor entwickelt. Der Autor war bei der Publikation federführend. Die Arbeit wurde vom Autor bei der Online-Konferenz präsentiert.

► Dang, Aslan, Seiderer und André (2020)

Der genutzte Systemprototyp wurde in (Seiderer, Aslan et al., 2019) maßgeblich vom Autor dieser Arbeit entwickelt. Der Autor war zusätzlich wesentlich an der für die Untersuchung benötigten Software, dem Konzept der Wizard-of-Oz-Studie und der Publikation beteiligt.

► Ritschel, Seiderer und André (2020)

Der Autor war unterstützend an der Publikation beteiligt.

► Aslan, Seiderer, Dang, Rädler und André (2020a, 2020b)

Der Autor war maßgeblich an der Planung, dem Studienaufbau und der Publikationen beteiligt. Die Hardware und Software (inklusive der Herzratenberechnung) der PiHeart-Haupteinheit und -Ausgabeeinheit (ausgenommen der Silikonhülle und Mechanik) wurden ausschließlich vom Autor entwickelt.

A.2 Durchgeführte Reviews

Jahr	Art	Anzahl	Name
2016	Journal	2	Hindawi: Advances in Human-Computer Interaction
2017	Konferenz	1	ACM IUI
2018	Workshop	2	ACM ICMI MHFI Workshop
2019	Journal	1	IEEE Transactions on Affective Computing
2019	Journal	2	ACM Transactions on Internet of Things
2020	Konferenz	1	ACM ICMI
2021	Journal	1	IEEE Transactions on Affective Computing

B Lehre

B.1 Betreute Lehrveranstaltungen

Art ¹	Lehrveranstaltung	Semester ²
M Seminar	Seminar: Opportunistic Human Activity and Context Recognition	S '15
M Vorlesung	Multimedia I: Usability Engineering	W '15 - '17
M Vorlesung	Interaction Design and Engineering (Elite)	W '20
M Praktikum	Multimodale Echtzeitsignalverarbeitung (verschiedene Themen)	S '16 - '18
M Praktikum	Praktikum Usability Engineering (verschiedene Themen)	S '14 - '17
M Praktikum	Projektmodul: Integration von Filtern / Features für Biosignalanalyse in SSI	S '15
M Praktikum	Projektmodul: Arduinobasiertes Wearable mit Biosensoren	S '16
M Praktikum	Projektmodul: Integration des Axiivity WAX-9-Sensors in SSI	S '16
M Praktikum	Projektmodul: Audioklassifizierung mit Neuronalen Netzen auf Android-Smartphones (mobileSSI)	S '17
M Praktikum	Projektmodul: Automatische Erstellung Neuronaler Netze für die Audioklassifizierung	S '17
M Praktikum	Projektmodul: Apache Spark: Anwendungen in der Echtzeitsignalverarbeitung und zur Visualisierung	W '18
B Praktikum	Praxismodul: Hierarchische Audioaktivitätserkennung mit SSI	W '15
B Praktikum	Multimedia Projekt (verschiedene Themen)	W '17 - '20
B Praktikum	Praktikum Interaction Design and Engineering for Health Care Applications	W '20
B Vorlesung	Multimedia Grundlagen 2 / Grundlagen der HCI	S '19 - '21

B.2 Betreute Abschlussarbeiten

Jahr	Art ¹	Titel	Name
2015	B	Recognition of Activities from Wearable Sensors	Mohamed Diab
2017	B	Implementierung einer grafischen Benutzeroberfläche für ein Optimierungswerkzeug im maritimen Bereich	Alexey Kvindt
2017	M	Gamifizierte Multi-Device-Interaktionen für Senioren-Übungen	Sebastian Klapprott
2018	B	Mobile Bilderkennung von Früchten unter Verwendung eines Deep Convolutional Neural Network	Marcel Henrich
2018	B	Konzeption und Realisierung eines Gas-Sensorik-Prototypen für das Smart Home	Christian Mittring
2018	M	Soundscape-Analyse durch Kombination tiefer neuronaler Netze zur Audio-Klassifizierung von kontinuierlichen und einzelnen Schallereignissen	Tobias Watzka
2018	M	Systematische Entwicklung und Evaluation eines Verfahrens für die automatische Generierung von tiefen neuronalen Netzen zur Audio-Klassifizierung, basierend auf evolutionären Algorithmen	Christian Rolet-scheck
2018	M	Nutzererlebnis einer greifbaren Herznachbildung zur Pulsdarstellung - Konzeption und Durchführung einer Feldstudie	Simon Rädler
2019	M	Untersuchung verschiedener Verfahren zur Kompensation und Kalibrierung von Gassensorik	Alex Barth
2019	B	Entwicklung einer ambienten Anzeige zur Überwachung der Luftqualität für einen Besprechungsraum	Andrés Caro Quintul

¹B = Bachelor, M = Master

²S = Sommersemester, W = Wintersemester

C (mobile)SSI Plugins

Folgende C++ Erweiterungen für das Open-Source-Framework (mobile)SSI wurden im Rahmen der vorliegenden Arbeit für verschiedene Projekte entwickelt oder mitentwickelt (Beteiligung). Die Links in der PDF-Version verweisen teilweise auf den Code bei Github. In der Regel stehen die Autoren in den entsprechenden C++-Headern.

Systeme sind folgendermaßen abgekürzt: W = Windows, L = Linux, A = Android. Nicht getestete Plattformen wurden nicht angegeben.

Art	Name	Kurzbeschreibung	System
Protokolle	mqtt	MQTT-Client	WL
	xmpp	XMPP-Client	WL
	websocket	WebSocket- und HTTP-Server	WLA
Sensoren	sensingtex	SensingTex Druckmatten	W
	touchmouse	Microsoft Touchmouse	W
	wax9	WAX 9 (Beteiligung)	W
	generic_serial	Daten von Geräten mit serieller Schnittstelle	WL
	rawinput	HID-Geräte ¹ : Maus, Keyboard, Controller	W
	androidSensors	Android-Sensor-API ² (Beteiligung)	A
	winsensors	Windows-Sensor-API ³	W
Features	pressuremat	Druckmatten-Features z. B. Center of Pressure (Beteiligung)	WLA
	praat	Praat ⁴ -Audiofeatures (Beteiligung)	W
Filter	libsox	SoX ⁵ -Audiofilter	W
Sonstige	eventinterpreter	Verarbeitung von SSI-Events mit Lua-Skripten	W
	fastDTW	fastDTW-Algorithmus ⁶	WLA
	FileBinaryLZ4	LZ4-Datenkompression ⁷ für SSI-Streams	WLA

¹<https://docs.microsoft.com/en-us/windows/win32/inputdev/about-raw-input> (27.02.2020)

²https://developer.android.com/guide/topics/sensors/sensors_overview (27.02.2020)

³<https://docs.microsoft.com/en-us/windows/win32/sensorsapi/portal> (27.02.2020)

⁴<http://www.fon.hum.uva.nl/praat/> (27.02.2020)

⁵<http://sox.sourceforge.net/> (27.02.2020)

⁶<https://github.com/melode11/FastDTW-x> (27.02.2020)

⁷<https://lz4.github.io/lz4/> (27.02.2020)

Literatur

- Ablart, D., Frier, W., Limerick, H., Georgiou, O. & Obrist, M. (2019). Using Ultrasonic Mid-air Haptic Patterns in Multi-Modal User Experiences. *2019 IEEE International Symposium on Haptic, Audio and Visual Environments and Games (HAVE)*, 1–6 (siehe S. 104).
- Adadi, A. & Berrada, M. (2018). Peeking inside the black-box: A survey on Explainable Artificial Intelligence (XAI). *IEEE Access*, 6, 52138–52160 (siehe S. 260).
- Adevi, A. A. & Mårtensson, F. (2013). Stress rehabilitation through garden therapy: The garden as a place in the recovery from stress. *Urban forestry & urban greening*, 12(2), 230–237 (siehe S. 113).
- Agarwal, A. & Zesch, T. (2019). German End-to-end Speech Recognition based on DeepSpeech. *Proceedings of the 15th Conference on Natural Language Processing (KONVENS 2019): Long Papers*, 111–119 (siehe S. 192).
- Air quality guidelines: global update 2005: particulate matter, ozone, nitrogen dioxide, and sulfur dioxide*. (2006). World Health Organization. (Siehe S. 97).
- Akbar, F., Mark, G., Pavlidis, I. & Gutierrez-Osuna, R. (2019). An empirical study comparing unobtrusive physiological sensors for stress detection in computer work. *Sensors*, 19(17), 3766 (siehe S. 66, 67, 70, 71, 78).
- Aked, J., Marks, N., Cordon, C. & Thompson, S. (2008). Five ways to well-being: The evidence. *London: new economics foundation* (siehe S. 14).
- Al Osman, H., Dong, H. & El Saddik, A. (2016). Ubiquitous biofeedback serious game for stress management. *IEEE Access*, 4, 1274–1286 (siehe S. 84).
- Aletta, F., Oberman, T. & Kang, J. (2018). Associations between positive health-related effects and soundscapes perceptual constructs: A systematic review. *International journal of environmental research and public health*, 15(11), 2392 (siehe S. 102).
- Aletta, F. & Xiao, J. (2018). What are the current priorities and challenges for (urban) soundscape research? *Challenges*, 9(1), 16 (siehe S. 101).
- Allen, J. G., MacNaughton, P., Cedeno-Laurent, J. G., Cao, X., Flanagan, S., Vallarino, J., Rueda, F., Donnelly-McLay, D. & Spengler, J. D. (2018). Airplane pilot flight performance on 21 maneuvers in a flight simulator under varying carbon dioxide concentrations. *Journal of exposure science & environmental epidemiology*, 1 (siehe S. 98).
- Allen, J. G., MacNaughton, P., Satish, U., Santanam, S., Vallarino, J. & Spengler, J. D. (2016). Associations of Cognitive Function Scores with Carbon Dioxide, Ventilation, and Volatile Organic Compound Exposures in Office Workers: A Controlled Exposure Study of Green and Conventional Office Environments. *Environmental Health Perspectives*, 124(6), 805 (siehe S. 97, 98).
- Altarriba Bertran, F., Jhaveri, S., Lutz, R., Isbister, K. & Wilde, D. (2019). Making sense of human-food interaction. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–13 (siehe S. 141).
- Amershi, S., Cakmak, M., Knox, W. B. & Kulesza, T. (2014). Power to the people: The role of humans in interactive machine learning. *Ai Magazine*, 35(4), 105–120 (siehe S. 178).
- Amft, O., Junker, H. & Tröster, G. (2005). Detection of eating and drinking arm gestures using inertial body-worn sensors. *Ninth IEEE International Symposium on Wearable Computers (ISWC'05)*, 160–163 (siehe S. 143).
- Amft, O., Stäger, M., Lukowicz, P. & Tröster, G. (2005). Analysis of chewing sounds for dietary monitoring. *International Conference on Ubiquitous Computing*, 56–72 (siehe S. 143).
- Amiriparian, S., Gerczuk, M., Ottl, S., Cummins, N., Freitag, M., Pugachevskiy, S., Baird, A. & Schuller, B. W. (2017). Snore Sound Classification Using Image-Based Deep Spectrum Features. *INTERSPEECH*, 3512–3516 (siehe S. 61).
- Angara, P., Jiménez, M., Agarwal, K., Jain, H., Jain, R., Stege, U., Ganti, S., Müller, H. A. & Ng, J. W. (2017). Foodie fooderson a conversational agent for the smart kitchen. *Proceedings of the 27th Annual International Conference on Computer Science and Software Engineering*, 247–253 (siehe S. 190).
- Angelini, L., Caon, M., Carrino, S., Bergeron, L., Nyffeler, N., Jean-Mairet, M. & Mugellini, E. (2013). Designing a desirable smart bracelet for older adults. *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*, 425–434 (siehe S. 206).
- Angelini, L., Mugellini, E., Abou Khaled, O., Röcke, C., Guye, S., Porcelli, S., Mastropietro, A., Rizzo, G., Boqué, N., Subias, P. et al. (2019). The NESTORE e-coach. *ICPS Proceedings, PETRA'19: Proceedings of the 12th ACM International Conference on Pervasive Technologies Related to Assistive Environments*, 5–7 June 2019, Rhodes, Greece, (ARTICLE) (siehe S. 190, 192, 248).
- Arnrich, B., Setz, C., La Marca, R., Tröster, G. & Ehlert, U. (2009). What does your chair know about your stress level? *IEEE Transactions on Information Technology in Biomedicine*, 14(2), 207–214 (siehe S. 71, 79).

- Årsand, E., Muzny, M., Bradway, M., Muzik, J. & Hartvigsen, G. (2015). Performance of the first combined smartwatch and smartphone diabetes diary application study. *Journal of diabetes science and technology*, 9(3), 556–563 (siehe S. 162).
- Aslan, I., Burkhardt, H., Kraus, J. & André, E. (2016). Hold my heart and breathe with me: Tangible somaesthetic designs. *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*, 1–6 (siehe S. 85).
- Aslan, I., Seiderer, A., Dang, C. T., Rädler, S. & André, E. (2020a). PiHearts: Resonating Experiences of Self and Others Enabled by a Tangible Somaesthetic Design. *Proceedings of the 2020 International Conference on Multimodal Interaction*, 433–441. <https://doi.org/10.1145/3382507.3418848> (siehe S. 83, 261, 266)
- Aslan, I., Seiderer, A., Dang, C. T., Rädler, S. & André, E. (2020b). Resonating experiences of self and others enabled by a tangible somaesthetic design. *arXiv preprint arXiv:2005.02304* (siehe S. 83, 261, 266).
- Azuma, K., Kagi, N., Yanagi, U. & Osawa, H. (2018). Effects of low-level inhalation exposure to carbon dioxide in indoor environments: A short review on human health and psychomotor performance. *Environment International*, 121, 51–56 (siehe S. 97).
- Bäck, T., Eiben, A. E. & van der Vaart, N. A. (2000). An empirical study on GAs “without parameters”. *International Conference on Parallel Problem Solving from Nature*, 315–324 (siehe S. 108, 109).
- Bäck, T. (1992). The interaction of mutation rate, selection, and self-adaptation within a genetic algorithm. *Proc. 2nd Conference of Parallel Problem Solving from Nature*, 1992 (siehe S. 109).
- Baltrušaitis, T., Robinson, P. & Morency, L.-P. (2016). Openface: an open source facial behavior analysis toolkit. *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 1–10 (siehe S. 40).
- Bao, L. & Intille, S. S. (2004). Activity recognition from user-annotated acceleration data. *International conference on pervasive computing*, 1–17 (siehe S. 178).
- Barraquand, R. & Crowley, J. L. (2008). Learning polite behavior with situation models. *2008 3rd ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 209–216 (siehe S. 182).
- Barsan, N. & Weimar, U. (2003). Understanding the fundamental principles of metal oxide based gas sensors; the example of CO sensing with SnO₂ sensors in the presence of humidity. *Journal of Physics: Condensed Matter*, 15(20), R813 (siehe S. 33).
- Bartl, A., Bosch, S., Brandt, M., Dittrich, M. & Lugin, B. (2016). The Influence of a social robot’s persona on how it is perceived and accepted by elderly users. *International Conference on Social Robotics*, 681–691 (siehe S. 243, 244).
- Baur, T., Mehlmann, G., Damian, I., Lingens, F., Wagner, J., Lugin, B., André, E. & Gebhard, P. (2015). Context-Aware Automated Analysis and Annotation of Social Human-Agent Interactions. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 5(2), 1–33 (siehe S. 175).
- Beck, C., Straub, A., Breitner, S., Cyrys, J., Philipp, A., Rathmann, J., Schneider, A., Wolf, K. & Jacobeit, J. (2018). Air temperature characteristics of local climate zones in the Augsburg urban area (Bavaria, southern Germany) under varying synoptic conditions. *Urban Climate*, 25, 152–166 (siehe S. 115).
- Begum, M. N. & Johnson, C. S. (2010). A review of the literature on dehydration in the institutionalized elderly. *e-SPEN, the European e-Journal of Clinical Nutrition and Metabolism*, 5(1), e47–e53 (siehe S. 209).
- Beigl, M., Gellersen, H.-W. & Schmidt, A. (2001). Mediapups: experience with design and use of computer-augmented everyday artefacts. *Computer Networks*, 35(4), 401–409 (siehe S. 143).
- Bentley, F., Tollmar, K., Stephenson, P., Levy, L., Jones, B., Robertson, S., Price, E., Catrambone, R. & Wilson, J. (2013). Health Mashups: Presenting statistical patterns between wellbeing data and context in natural language to promote behavior change. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 20(5), 30 (siehe S. 205).
- Bhattacharya, S. & Lane, N. D. (2016). From smart to deep: Robust activity recognition on smartwatches using deep learning. *2016 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*, 1–6 (siehe S. 174).
- Bifet, A., Holmes, G., Kirkby, R. & Pfahringer, B. (2010). Moa: Massive online analysis. *Journal of Machine Learning Research*, 11(May), 1601–1604 (siehe S. 178).
- Bittner, B., Aslan, I., Dang, C. T. & André, E. (2019). Of smarthomes, IoT plants, and implicit interaction design. *Proceedings of the Thirteenth International Conference on Tangible, Embedded, and Embodied Interaction*, 145–154 (siehe S. 113).
- Bocklisch, T., Faulkner, J., Pawlowski, N. & Nichol, A. (2017). Rasa: Open source language understanding and dialogue management. *arXiv preprint arXiv:1712.05181* (siehe S. 192).
- Bosch, S. (2016). Unerwartete Kohlenmonoxid-Gefahren: Tischgrills, Wasserpfeifen und Holzpellets. *Biologie in unserer Zeit*, 46(2), 114–117 (siehe S. 98).
- Brand, P., Bertram, J., Chaker, A., Jörres, R., Kronseder, A., Kraus, T. & Gube, M. (2016). Biological effects of inhaled nitrogen dioxide in healthy human subjects. *International archives of occupational and environmental health*, 89(6), 1017–1024 (siehe S. 98).
- Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2), 77–101 (siehe S. 90).
- Brooke, J. (1996). SUS - A quick and dirty usability scale (P. Jordan, B. Thomas, B. A. Weerdmeester & A. McClelland, Hrsg.). *Usability evaluation in industry*, 189(194), 4–7 (siehe S. 246).
- Budde, M., Schwarz, A. D., Müller, T., Laquai, B., Streibl, N., Schindler, G., Köpke, M., Riedel, T., Dittler, A. & Beigl, M. (2018). Potential and limitations of the low-cost SDS011 particle sensor for monitoring urban air quality. *3rd International Conference on Atmospheric Dust*, 6–12 (siehe S. 26).
- Butz, A. & Krüger, A. (2017). *Mensch-Maschine-Interaktion*. Walter de Gruyter GmbH & Co KG. (Siehe S. 128).

- Buxton, B. (1995). Integrating the Periphery & Context: A New Model of Telematics. *Graphics Interface*, 239–239 (siehe S. 103).
- Buxton, R. T., Pearson, A. L., Allou, C., Fristrup, K. & Wittemyer, G. (2021). A synthesis of health benefits of natural sounds and their distribution in national parks. *Proceedings of the National Academy of Sciences*, 118(14) (siehe S. 102, 118).
- Calibo, T. K., Blanco, J. A. & Firebaugh, S. L. (2013). Cognitive stress recognition. *2013 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, 1471–1475 (siehe S. 70).
- Carter, M. C., Burley, V. J., Nykjaer, C. & Cade, J. E. (2013). Adherence to a smartphone application for weight loss compared to website and paper diary: pilot randomized controlled trial. *Journal of medical Internet research*, 15(4), e32 (siehe S. 161).
- Casas, J., Mugellini, E. & Khaled, O. A. (2018). Food diary coaching chatbot. *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*, 1676–1680 (siehe S. 190).
- Chan, F. K. S., Griffiths, J. A., Higgitt, D., Xu, S., Zhu, F., Tang, Y.-T., Xu, Y. & Thorne, C. R. (2018). “Sponge City” in China—a breakthrough of planning and flood risk management in the urban context. *Land Use Policy*, 76, 772–778 (siehe S. 113).
- Chang, C.-C. & Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM transactions on intelligent systems and technology (TIST)*, 2(3), 1–27 (siehe S. 148).
- Chen, Z., Peng, W., Peiris, R., Chan, L. & Minamizawa, K. (2017). Thermoreality: thermally enriched head mounted displays. *SIGGRAPH Asia 2017 Emerging Technologies* (S. 1–2). (Siehe S. 104).
- Cheng, A., Raghavaraju, V., Kanugo, J., Handrianto, Y. P. & Shang, Y. (2018). Development and evaluation of a healthy coping voice interface application using the Google home for elderly patients with type 2 diabetes. *2018 15th IEEE Annual Consumer Communications & Networking Conference (CCNC)*, 1–5 (siehe S. 190).
- Consolvo, S., McDonald, D. W., Toscos, T., Chen, M. Y., Froehlich, J., Harrison, B., Klasnja, P., LaMarca, A., LeGrand, L., Libby, R. et al. (2008). Activity sensing in the wild: a field trial of ubifit garden. *Proceedings of the SIGCHI conference on human factors in computing systems*, 1797–1806 (siehe S. 205).
- Consolvo, S., Roessler, P. & Shelton, B. E. (2004). The CareNet display: lessons learned from an in home evaluation of an ambient display. *International conference on ubiquitous computing*, 1–17 (siehe S. 206).
- Consolvo, S. & Towle, J. (2005). Evaluating an ambient display for the home. *CHI’05 extended abstracts*, 1304–1307 (siehe S. 130, 224).
- Cordeiro, F., Bales, E., Cherry, E. & Fogarty, J. (2015). Rethinking the mobile food journal: Exploring opportunities for lightweight photo-based capture. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 3207–3216 (siehe S. 161).
- Cordeiro, F., Epstein, D. A., Thomaz, E., Bales, E., Jagannathan, A. K., Abowd, G. D. & Fogarty, J. (2015). Barriers and negative nudges: Exploring challenges in food journaling. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 1159–1162 (siehe S. 160, 163, 164).
- Cornejo, R., Tentori, M. & Favela, J. (2013). Enriching in-person encounters through social media: A study on family connectedness for the elderly. *International Journal of Human-Computer Studies*, 71(9), 889–899 (siehe S. 206, 225).
- Cortes, C. & Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3), 273–297 (siehe S. 45).
- Cosentino, S., Sessa, S. & Takanishi, A. (2016). Quantitative laughter detection, measurement, and classification—A critical survey. *IEEE Reviews in Biomedical engineering*, 9, 148–162 (siehe S. 68).
- Costa, J., Adams, A. T., Jung, M. F., Guimbretiére, F. & Choudhury, T. (2016). EmotionCheck: leveraging bodily signals and false feedback to regulate our emotions. *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 758–769 (siehe S. 84).
- Coucke, A., Saade, A., Ball, A., Bluche, T., Caulier, A., Leroy, D., Doumouro, C., Gisselbrecht, T., Caltagirone, F., Lavril, T. et al. (2018). Snips voice platform: an embedded spoken language understanding system for private-by-design voice interfaces. *arXiv preprint arXiv:1805.10190* (siehe S. 190, 191, 196).
- Cruz-Maya, A. & Tapus, A. (2016). Teaching nutrition and healthy eating by using multimedia with a Kompai robot: Effects of stress and user’s personality. *2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)*, 644–649 (siehe S. 182).
- Dacremont, C. (1995). Spectral composition of eating sounds generated by crispy, crunchy and crackly foods. *Journal of texture studies*, 26(1), 27–43 (siehe S. 148).
- Damian, I., Dietz, M. & André, E. (2018). The SSJ Framework: Augmenting Social Interactions Using Mobile Signal Processing and Live Feedback. *Frontiers in ICT*, 5, 13 (siehe S. 41).
- Dang, C. T., Aslan, I., Seiderer, A. & André, E. (2020). Exploring Ambient Air Quality Notifications for Smart Rooms, 102–107. <https://doi.org/10.1145/3411170.3411229> (siehe S. 263, 266)
- Dang, C. T., Seiderer, A. & André, E. (2018). Theodor: A Step Towards Smart Home Applications with Electronic Noses. *Proceedings of the 5th international Workshop on Sensor-based Activity Recognition and Interaction*, 11. <https://doi.org/10.1145/3266157.3266215> (siehe S. 37, 152, 262, 265)
- Dayarathna, M. & Perera, S. (2018). Recent advancements in event processing. *ACM Computing Surveys (CSUR)*, 51(2), 1–36 (siehe S. 42).
- De Schutter, B. & Vanden Abeele, V. (2010). Designing meaningful play within the psycho-social context of older adults. *Proceedings of the 3rd International Conference on Fun and Games*, 84–93 (siehe S. 237).

- Demerouti, E., Bakker, A. B., Nachreiner, F. & Schaufeli, W. B. (2001). The job demands-resources model of burnout. *Journal of Applied psychology*, 86(3), 499 (siehe S. 70).
- Deterding, S., Dixon, D., Khaled, R. & Nacke, L. (2011). From Game Design Elements to Gamefulness: Defining "Gamification". *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, 9–15. <https://doi.org/10.1145/2181037.2181040> (siehe S. 236)
- Di Stasi, L. L., Catena, A., Canas, J. J., Macknik, S. L. & Martinez-Conde, S. (2013). Saccadic velocity as an arousal index in naturalistic tasks. *Neuroscience & Biobehavioral Reviews*, 37(5), 968–975 (siehe S. 71).
- Ding, R., Li, X., Nie, L., Li, J., Si, X., Chu, D., Liu, G. & Zhan, D. (2019). Empirical study and improvement on deep transfer learning for human activity recognition. *Sensors*, 19(1), 57 (siehe S. 62).
- Domínguez-Jiménez, J., Campo-Landines, K., Martínez-Santos, J., Delahoz, E. & Contreras-Ortiz, S. (2020). A machine learning model for emotion recognition from physiological signals. *Biomedical Signal Processing and Control*, 55, 101646 (siehe S. 67).
- Dong, T., Churchill, E. F. & Nichols, J. (2016). Understanding the challenges of designing and developing multi-device experiences. *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*, 62–72 (siehe S. 160, 164).
- Efe, A. (2017). Using Olfactory Displays as a Nontraditional Interface in Human Computer Interaction. *Journal of learning and teaching in digital age*, 2(2), 14–25 (siehe S. 104).
- Eiben, A. & Smith, J. (2015). *Introduction to Evolutionary Computing*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-44874-8>. (Siehe S. 106, 107, 108, 109)
- Eyben, F., Weninger, F., Gross, F. & Schuller, B. (2013). Recent developments in opensmile, the munich open-source multimedia feature extractor. *Proceedings of the 21st ACM international conference on Multimedia*, 835–838 (siehe S. 148).
- Eyben, F., Wöllmer, M. & Schuller, B. (2010). Opensmile: the munich versatile and fast open-source audio feature extractor. *Proceedings of the 18th ACM international conference on Multimedia*, 1459–1462 (siehe S. 40, 148).
- Fails, J. A. & Olsen Jr, D. R. (2003). Interactive machine learning. *Proceedings of the 8th international conference on Intelligent user interfaces*, 39–45 (siehe S. 175).
- Faltaous, S., Haas, G., Barrios, L., **Seiderer, A.**, Rauh, S. F., Chae, H. J., Schneegass, S. & Alt, F. (2019). BrainShare: A Glimpse of Social Interaction for Locked-in Syndrome Patients. *CHI'19 Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3290607.3312754> (siehe S. 263, 265)
- Fan, R.-E., Chang, K.-W., Hsieh, C.-J., Wang, X.-R. & Lin, C.-J. (2008). LIBLINEAR: A library for large linear classification. *Journal of machine learning research*, 9(Aug), 1871–1874 (siehe S. 178).
- Fasola, J. & Matarić, M. J. (2013). A socially assistive robot exercise coach for the elderly. *Journal of Human-Robot Interaction*, 2(2), 3–32 (siehe S. 243).
- Fernando, C., Banarse, D., Reynolds, M., Besse, F., Pfau, D., Jaderberg, M., Lanctot, M. & Wierstra, D. (2016). Convolution by Evolution: Differentiable Pattern Producing Networks. *Proceedings of the Genetic and Evolutionary Computation Conference 2016*, 109–116. <https://doi.org/10.1145/2908812.2908890> (siehe S. 105)
- Fiebig, A. & Schulte-Fortkamp, B. (2019). Soundscape-Fortschritte in der Standardisierung auf internationaler Ebene. *Akustik Journal*, 01/19 (siehe S. 101).
- Fisk, A., Czaja, S., Rogers, W., Charness, N. & Sharit, J. (2009). *Designing for Older Adults: Principles and Creative Human Factors Approaches, Second Edition*. CRC Press. (Siehe S. 203).
- Flutura, S., **Seiderer, A.**, Aslan, I., Dang, C. T., Schwarz, R., Schiller, D. & André, E. (2018). DrinkWatch: A Mobile Wellbeing Application Based on Interactive and Cooperative Machine Learning. *Proceedings of the 2018 International Conference on Digital Health*, 65–74. <https://doi.org/10.1145/3194658.3194666> (siehe S. 174, 177, 262, 265)
- Flutura, S., **Seiderer, A.**, Aslan, I., Dietz, M., Schiller, D., Beck, C., Rathmann, J. & André, E. (2019). Mobile Sensing for Wellbeing Estimation of Urban Green Using Physiological Signals. *Proceedings of the 5th EAI International Conference on Smart Objects and Technologies for Social Good*, 249–254. <https://doi.org/10.1145/3342428.3342655> (siehe S. 114, 115, 117, 261, 266)
- Flutura, S., **Seiderer, A.**, Huber, T., Weitz, K., Aslan, I., Schlagowski, R., André, E. & Rathmann, J. (2020). Interactive Machine Learning and Explainability in Mobile Classification of Forest-Aesthetics. *Proceedings of the 6th EAI International Conference on Smart Objects and Technologies for Social Good*, 90–95. <https://doi.org/10.1145/3411170.3411225> (siehe S. 181, 263, 266)
- Flutura, S., Wagner, J., Lingensfelder, F., **Seiderer, A.** & André, E. (2016a). Laughter detection in the wild: demonstrating a tool for mobile social signal processing and visualization. *Proceedings of the 18th ACM International Conference on Multimodal Interaction - ICMI 2016*, 406–407. <https://doi.org/10.1145/2993148.2998524> (siehe S. 80, 261, 264)
- Flutura, S., Wagner, J., Lingensfelder, F., **Seiderer, A.** & André, E. (2016b). MobileSSI - A Multi-modal Framework for Social Signal Interpretation on Mobile Devices. *2016 12th International Conference on Intelligent Environments (IE)*, 210–213. <https://doi.org/10.1109/IE.2016.47> (siehe S. 79, 261, 264)
- Flutura, S., Wagner, J., Lingensfelder, F., **Seiderer, A.** & André, E. (2016c). MobileSSI: asynchronous fusion for social signal interpretation in the wild. *Proceedings of the 18th ACM International Conference on Multimodal Interaction - ICMI 2016*, 266–273. <https://doi.org/10.1145/2993148.2993164> (siehe S. 41, 80, 261, 264)
- Folstein, M. F., Robins, L. N. & Helzer, J. E. (1983). The mini-mental state examination. *Archives of general psychiatry*, 40(7), 812–812 (siehe S. 209).
- Forgy, C. L. (1989). Rete: A fast algorithm for the many pattern/many object pattern match problem. *Readings in Artificial Intelligence and Databases* (S. 547–559). Elsevier. (Siehe S. 231).

- Fortmann, J., Stratmann, T. C., Boll, S., Poppinga, B. & Heuten, W. (2013). Make me move at work! An ambient light display to increase physical activity. *Proceedings of the 7th international conference on pervasive computing technologies for healthcare*, 274–277 (siehe S. 130).
- Franco, R. Z., Fallaize, R., Lovegrove, J. A. & Hwang, F. (2016). Popular nutrition-related mobile apps: a feature assessment. *JMIR mHealth and uHealth*, 4(3), e85 (siehe S. 161, 163, 173, 174).
- Frank, E., Hall, M. A. & Witten, I. H. (2016). *Online Appendix for „Data Mining: Practical Machine Learning Tools and Techniques“* (Fourth Edition). Morgan Kaufmann. (Siehe S. 153).
- Frank, K., Robertson, P., Gross, M. & Wiesner, K. (2013). Sensor-based identification of human stress levels. *2013 IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*, 127–132 (siehe S. 70).
- Frenzel, A. C., Götz, T. & Pekrun, R. (2009). Emotionen. *Pädagogische Psychologie* (S. 205–231). Springer. (Siehe S. 117).
- Fukumoto, K., Terada, T. & Tsukamoto, M. (2013). A smile/laughter recognition mechanism for smile-based life logging. *Proceedings of the 4th Augmented Human International Conference*, 213–220 (siehe S. 82).
- Gaffin, J. M., Hauptman, M., Petty, C. R., Sheehan, W. J., Lai, P. S., Wolfson, J. M., Gold, D. R., Coull, B. A., Koutrakis, P. & Phipatanakul, W. (2018). Nitrogen dioxide exposure in school classrooms of inner-city children with asthma. *Journal of Allergy and Clinical Immunology*, 141(6), 2249–2255 (siehe S. 98).
- Gao, Y., Zhang, N., Wang, H., Ding, X., Ye, X., Chen, G. & Cao, Y. (2016). iHear food: eating detection using commodity bluetooth headsets. *2016 IEEE First International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*, 163–172 (siehe S. 143).
- Gatt, A. & Reiter, E. (2009). SimpleNLG: A realisation engine for practical applications. *Proceedings of the 12th European Workshop on Natural Language Generation (ENLG 2009)*, 90–93 (siehe S. 193).
- Gebhard, P., Mehlmann, G. & Kipp, M. (2012). Visual SceneMaker—a tool for authoring interactive virtual characters. *Journal on Multimodal User Interfaces*, 6(1-2), 3–11 (siehe S. 186).
- Gebhard, P., Schröder, M., Charfuelan, M., Endres, C., Kipp, M., Pammi, S., Rumpler, M. & Türk, O. (2008). IDEAS4Games: building expressive virtual characters for computer games. *International Workshop on Intelligent Virtual Agents*, 426–440 (siehe S. 243, 244).
- Gervais, R., Frey, J., Gay, A., Lotte, F. & Hachet, M. (2016). Tobe: Tangible out-of-body experience. *Proceedings of the TEI'16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction*, 227–235 (siehe S. 83).
- Gesundheitliche Bewertung von Kohlendioxid in der Innenraumluft. (2008). *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz*, 51(11), 1358–1369 (siehe S. 98, 128).
- Gillian, N. & Paradiso, J. A. (2014). The gesture recognition toolkit. *The Journal of Machine Learning Research*, 15(1), 3483–3487 (siehe S. 40, 238).
- Giraud, T., Soury, M., Hua, J., Delaborde, A., Tahon, M., Jauregui, D. A. G., Eyharabide, V., Filaire, E., Le Scanff, C., Devillers, L. et al. (2013). Multimodal expressions of stress during a public speaking task: Collection, annotation and global analyses. *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, 417–422 (siehe S. 71).
- Goodfellow, I., Bengio, Y. & Courville, A. (2016). *Deep Learning* [<http://www.deeplearningbook.org>]. MIT Press. (Siehe S. 45).
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. & Bengio, Y. (2014). Generative Adversarial Nets. In Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence & K. Q. Weinberger (Hrsg.), *Advances in Neural Information Processing Systems 27* (S. 2672–2680). Curran Associates, Inc. (Siehe S. 260).
- Gordon, G., Spaulding, S., Westlund, J. K., Lee, J. J., Plummer, L., Martinez, M., Das, M. & Breazeal, C. (2016). Affective personalization of a social robot tutor for children's second language skills. *Thirtieth AAAI Conference on Artificial Intelligence* (siehe S. 182).
- Gradl, S., Wirth, M., Richer, R., Rohleder, N. & Eskofier, B. M. (2019). An overview of the feasibility of permanent, real-time, unobtrusive stress measurement with current wearables. *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare*, 360–365 (siehe S. 67).
- Greene, S., Thapliyal, H. & Caban-Holt, A. (2016). A survey of affective computing for stress detection: Evaluating technologies in stress detection for better health. *IEEE Consumer Electronics Magazine*, 5(4), 44–56 (siehe S. 67).
- Guo, X., Polania, L. & Barner, K. (2018). Smile detection in the wild based on transfer learning. *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, 679–686 (siehe S. 62).
- Gustafsson, A. & Gyllenswärd, M. (2005). The Power-aware Cord: Energy Awareness Through Ambient Information Display. *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, 1423–1426. <https://doi.org/10.1145/1056808.1056932> (siehe S. 121).
- Hagen, K. (2010). Nutritional information: traffic light labelling is the best way to reach consumers. *Weekly Report*, 6(19), 141–151 (siehe S. 195).
- Hagerer, G., Cummins, N., Eyben, F. & Schuller, B. W. (2017). "Did you laugh enough today? Deep Neural Networks for Mobile and Wearable Laughter Trackers. *INTERSPEECH*, 2044–2045 (siehe S. 69).
- Hammer, S., Kirchner, K., André, E. & Lugin, B. (2017). Touch or Talk? Comparing Social Robots and Tablet PCs for an Elderly Assistant Recommender System. *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, 129–130 (siehe S. 182).

- Hammer, S., Lugrin, B., Bogomolov, S., Janowski, K. & André, E. (2016). Investigating politeness strategies and their persuasiveness for a robotic elderly assistant. *International Conference on Persuasive Technology*, 315–326 (siehe S. 188, 224, 243).
- Hammer, S., Seiderer, A., André, E., Rist, T., Kastrinaki, S., Hondrou, C., Raouzaïou, A., Karpouzis, K. & Kollias, S. (2015). Design of a lifestyle recommender system for the elderly: requirement gatherings in Germany and Greece. *Proceedings of the 8th ACM International Conference on Pervasive Technologies Related to Assistive Environments - PETRA '15*, 1–8. <https://doi.org/10.1145/2769493.2769559> (siehe S. 202, 204, 262, 263, 264)
- Hanst, P. L. (1986). IR-spectroscopy of the atmosphere. *Fresenius' Zeitschrift für analytische Chemie*, 324(6), 579–588 (siehe S. 31).
- Hantke, S., Weninger, F., Kurle, R., Ringeval, F., Batliner, A., Mousa, A. E.-D. & Schuller, B. (2016). I hear you eat and speak: Automatic recognition of eating condition and food type, use-cases, and impact on asr performance. *PloS one*, 11(5) (siehe S. 148).
- Harjuniemi, E., Colley, A., Rytlahti, P. & Häkkinen, J. (2020). IdleStripes shirt-wearable display of sedentary time. *Proceedings of the 9TH ACM International Symposium on Pervasive Displays*, 29–36 (siehe S. 104).
- Hart, S. G. & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in psychology* (S. 139–183). Elsevier. (Siehe S. 71).
- Haslam, C., Cruwys, T., Haslam, S. A. & Jetten, J. (2015). Social connectedness and health. *Encyclopaedia of geropsychology*, 2015, 46–1 (siehe S. 68).
- Hassenzahl, M., Burmester, M. & Koller, F. (2003). AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In G. Szwillus & J. Ziegler (Hrsg.), *Mensch & Computer 2003* (S. 187–196). Vieweg+Teubner Verlag. https://doi.org/10.1007/978-3-322-80058-9_19. (Siehe S. 246)
- Hatfield, E., Hatfield, C., Cacioppo, J. T. & Rapson, R. L. (1994). *Emotional Contagion*. Cambridge University Press. (Siehe S. 66).
- Hazlewood, W. R., Stolterman, E. & Connelly, K. (2011). Issues in Evaluating Ambient Displays in the Wild: Two Case Studies. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 877–886. <https://doi.org/10.1145/1978942.1979071> (siehe S. 121)
- Hering, E., Martin, R. & Stohrer, M. (2016). *Physik für Ingenieure* (12. Auflage). Springer Vieweg. <https://doi.org/10.1007/978-3-662-49355-7>. (Siehe S. 27, 28, 29, 99)
- Hering, E., Schönfelder, G. et al. (2012). *Sensoren in Wissenschaft und Technik* (1. Auflage). Vieweg+Teubner Verlag. <https://doi.org/10.1007/978-3-8348-8635-4>. (Siehe S. 25)
- Hernandez, J., Paredes, P., Roseway, A. & Czerwinski, M. (2014). Under pressure: sensing stress of computer users. *Proceedings of the SIGCHI conference on Human factors in computing systems*, 51–60 (siehe S. 71).
- Herpich, M., Rist, T., Seiderer, A. & André, E. (2017). Towards a Gamified Recommender System for the Elderly. *Proceedings of the 2017 International Conference on Digital Health - DH '17*, 211–215. <https://doi.org/10.1145/3079452.3079500> (siehe S. 202, 236, 262, 264)
- Hirano, S. H., Hayes, G. R. & Truong, K. N. (2015). uSmell: exploring the potential for gas sensors to classify odors in ubicomp applications relative to airflow and distance. *Personal and Ubiquitous Computing*, 19(1), 189–202 (siehe S. 143, 145, 152, 153, 154, 156).
- Holdsworth, M. A. (2019). Health, Wellness and Wellbeing. *Revue Interventions économiques. Papers in Political Economy*, (62) (siehe S. 13).
- Holzinger, A. (2016). Interactive machine learning for health informatics: when do we need the human-in-the-loop? *Brain Informatics*, 3(2), 119–131 (siehe S. 175).
- Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M. & Adam, H. (2017). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *CoRR*, abs/1704.04861 (siehe S. 56).
- Hoy, M. B. (2018). Alexa, Siri, Cortana, and more: an introduction to voice assistants. *Medical reference services quarterly*, 37(1), 81–88 (siehe S. 197).
- Hughes, J., O'Brien, J., Rodden, T., Rouncefield, M. & Viller, S. (2000). Patterns of home life: Informing design for domestic environments. *Personal Technologies*, 4(1), 25–38. <https://doi.org/10.1007/BF01613596> (siehe S. 122)
- Huijnen, C., Badii, A., van den Heuvel, H., Caleb-Solly, P. & Thiemert, D. (2011). “Maybe it becomes a buddy, but do not call it a robot” –Seamless cooperation between companion robotics and smart homes. *International Joint Conference on Ambient Intelligence*, 324–329 (siehe S. 243).
- Huynh, T. & Schiele, B. (2005). Analyzing features for activity recognition. *Proceedings of the 2005 joint conference on Smart objects and ambient intelligence: innovative context-aware services: usages and technologies*, 159–163 (siehe S. 178).
- IJsselstein, W., De Kort, Y. & Poels, K. (2013). The game experience questionnaire. *Eindhoven: Technische Universiteit Eindhoven*, 3–9 (siehe S. 90).
- Ionescu, R., Vancu, A., Moise, C. & Tomescu, A. (1999). Role of water vapour in the interaction of SnO₂ gas sensors with CO and CH₄. *Sensors and Actuators B: Chemical*, 61(1-3), 39–42 (siehe S. 33).
- Jackson, M. D. & McClelland, J. L. (1975). Sensory and cognitive determinants of reading speed. *Journal of Verbal Learning and Verbal Behavior*, 14(6), 565–574 (siehe S. 74).
- Jansen, M.-p., Heylen, D. K., Truong, K. P., Englebienne, G. & Nazareth, D. S. (2018). The MULAI Corpus: Multimodal Recordings of Spontaneous Laughter in Dyadic Interaction. *Proceedings of Laughter Workshop*, 58–63 (siehe S. 80).
- Jayakody, D. M., Friedland, P. L., Martins, R. N. & Sohrabi, H. R. (2018). Impact of aging on the auditory system and related cognitive functions: a narrative review. *Frontiers in neuroscience*, 12, 125 (siehe S. 203).

- Jennings, V., Browning, M. H. E. M. & Rigolon, A. (2019). Friend or Foe? An Overview of the Services and Disservices from Urban Green Spaces. *Urban Green Spaces: Public Health and Sustainability in the United States* (S. 7–30). Springer International Publishing. https://doi.org/10.1007/978-3-030-10469-6_2. (Siehe S. 95, 100)
- Jensen, W., Löchtefeld, M. & Knoche, H. (2019). ShadowLamp: An Ambient Display with Controllable Shadow Projection using Electrochromic Materials. *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–6 (siehe S. 104).
- Jesús-Azabal, M., Rojo, J., Moguel, E., Flores-Martin, D., Berrocal, J., García-Alonso, J. & Murillo, J. M. (2019). Voice Assistant to Remind Pharmacologic Treatment in Elders. *International Workshop on Gerontechnology*, 123–133 (siehe S. 190).
- Jia, W., Chen, H.-C., Yue, Y., Li, Z., Fernstrom, J., Bai, Y., Li, C. & Sun, M. (2014). Accuracy of food portion size estimation from digital pictures acquired by a chest-worn camera. *Public health nutrition*, 17(8), 1671–1681 (siehe S. 144).
- Johnson, D., Deterding, S., Kuhn, K.-A., Staneva, A., Stoyanov, S. & Hides, L. (2016). Gamification for health and wellbeing: A systematic review of the literature. *Internet interventions*, 6, 89–106 (siehe S. 236).
- Johnson, W. L., Mayer, R. E., André, E. & Rehm, M. (2005). Cross-Cultural Evaluation of Politeness in Tactics for Pedagogical Agents. *AIED*, 298–305 (siehe S. 244).
- Jones, T., Kay, D., Upton, P. & Upton, D. (2013). An evaluation of older adults use of iPads in eleven UK care-homes. *International Journal of Mobile Human Computer Interaction (IJMHCI)*, 5(3), 62–76 (siehe S. 206).
- Jozefowicz, R., Zaremba, W. & Sutskever, I. (2015). An empirical exploration of recurrent network architectures. *International Conference on Machine Learning*, 2342–2350 (siehe S. 105).
- Ju, W. & Leifer, L. (2008). The design of implicit interactions: Making interactive systems less obnoxious. *Design Issues*, 24(3), 72–84 (siehe S. 5, 103, 104).
- Julia, C., Hercberg, S., World Health Organization et al. (2017). Development of a new front-of-pack nutrition label in France: the five-colour Nutri-Score. *Public Health Panorama*, 3(04), 712–725 (siehe S. 190).
- Kabat-Zinn, J. (1982). An outpatient program in behavioral medicine for chronic pain patients based on the practice of mindfulness meditation: Theoretical considerations and preliminary results. *General hospital psychiatry*, 4(1), 33–47 (siehe S. 83).
- Kadomura, A., Li, C.-Y., Chen, Y.-C., Tsukada, K., Siio, I. & Chu, H.-h. (2013). Sensing fork: eating behavior detection utensil and mobile persuasive game. *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, 1551–1556 (siehe S. 143).
- Kalantarian, H. & Sarrafzadeh, M. (2015). Audio-based detection and evaluation of eating behavior using the smartwatch platform. *Computers in biology and medicine*, 65, 1–9 (siehe S. 143).
- Kalimeri, K. & Saitis, C. (2016). Exploring multimodal biosignal features for stress detection during indoor mobility. *Proceedings of the 18th ACM international conference on multimodal interaction*, 53–60 (siehe S. 113).
- Kaneda, T., Lee, M. & Pollard, K. (2011). SCL/PRB index of well-being in older populations. *Washington DC: Population Reference Bureau* (siehe S. 13).
- Karrer, K., Glaser, C., Clemens, C. & Bruder, C. (2009). Technikaffinität erfassen – der Fragebogen TA-EG. *ZMMS Spektrum*, 29 (siehe S. 246).
- Kemker, R., McClure, M., Abitino, A., Hayes, T. L. & Kanan, C. (2018). Measuring catastrophic forgetting in neural networks. *Thirty-second AAAI conference on artificial intelligence* (siehe S. 62).
- Kidd, C. D. & Breazeal, C. (2007). A robotic weight loss coach. *Proceedings of the national conference on artificial intelligence*, 22(2), 1985 (siehe S. 182).
- Kidd, C. D. & Breazeal, C. (2008). Robots at home: Understanding long-term human-robot interaction. *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 3230–3235 (siehe S. 182).
- Kientz, J. A., Choe, E. K., Birch, B., Maharaj, R., Fonville, A., Glasson, C. & Mundt, J. (2010). Heuristic evaluation of persuasive health technologies. *Proceedings of the 1st ACM International Health Informatics Symposium*, 555–564 (siehe S. 223).
- Kim, J. (2007). Bimodal emotion recognition using speech and physiological changes. *Robust speech recognition and understanding*, 265, 280 (siehe S. 75).
- Kim, K.-H., Kabir, E. & Jahan, S. A. (2018). Airborne bioaerosols and their impact on human health. *Journal of Environmental Sciences*, 67, 23–35 (siehe S. 100).
- Kim, K. H., Bang, S. W. & Kim, S. R. (2004). Emotion recognition system using short-term monitoring of physiological signals. *Medical and biological engineering and computing*, 42(3), 419–427 (siehe S. 113).
- Kim, M. & Rigazio, L. (2015). Deep Clustered Convolutional Kernels. *CoRR*, abs/1503.01824 (siehe S. 106).
- Kim, Y., Ji, S., Lee, H., Kim, J.-W., Yoo, S. & Lee, J. (2016). My doctor is keeping an eye on me!: Exploring the clinical applicability of a mobile food logger. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 5620–5631 (siehe S. 160, 161, 164, 173).
- Kirschbaum, C. (2008). *Biopsychologie von A bis Z*. Springer-Verlag. (Siehe S. 66).
- Knuth, D. E. (1985). *The Art of Computer Programming, Volume 2 (2nd Ed.): Seminumerical Algorithms*. Addison-Wesley Longman Publishing Co., Inc. (Siehe S. 45).
- Kobayashi, M., Kosugi, A., Takagi, H., Nemoto, M., Nemoto, K., Arai, T. & Yamada, Y. (2019). Effects of Age-Related Cognitive Decline on Elderly User Interactions with Voice-Based Dialogue Systems. *IFIP Conference on Human-Computer Interaction*, 53–74 (siehe S. 197).

- Koldijk, S., Neerincx, M. A. & Kraaij, W. (2018). Detecting Work Stress in Offices by Combining Unobtrusive Sensors. *IEEE Transactions on Affective Computing*, (2), 227–239 (siehe S. 67).
- Koldijk, S., Sappelli, M., Verberne, S., Neerincx, M. A. & Kraaij, W. (2014). The swell knowledge work dataset for stress and user modeling research. *Proceedings of the 16th international conference on multimodal interaction*, 291–298 (siehe S. 75).
- Kotera, Y., Richardson, M. & Sheffield, D. (2020). Effects of shinrin-yoku (forest bathing) and nature therapy on mental health: A systematic review and meta-analysis. *International Journal of Mental Health and Addiction*, 1–25 (siehe S. 113).
- Krafczyk, S., Schlamp, V., Dieterich, M., Haberhauer, P. & Brandt, T. (1999). Increased body sway at 3.5–8 Hz in patients with phobic postural vertigo. *Neuroscience letters*, 259(3), 149–152 (siehe S. 74).
- Kroos, C. & Plumbley, M. (2017). Neuroevolution for sound event detection in real life audio: A pilot study. In T. Virtanen, A. Mesaros, T. Heittola, A. Diment, E. Vincent, E. Benetos & B. Elizalde (Hrsg.), *DCASE 2017*. Tampere University of Technology. (Siehe S. 106).
- Kruijff-Korbayová, I., Oleari, E., Baroni, I., Kiefer, B., Zelati, M. C., Pozzi, C. & Sanna, A. (2014). Effects of off-activity talk in human-robot interaction with diabetic children. *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*, 649–654 (siehe S. 182).
- Lane, N. D., Lin, M., Mohammad, M., Yang, X., Lu, H., Cardone, G., Ali, S., Doryab, A., Berke, E., Campbell, A. T. et al. (2014). Bewell: Sensing sleep, physical activities and social interactions to promote wellbeing. *Mobile Networks and Applications*, 19(3), 345–359 (siehe S. 3, 205).
- Lang, P. J., Bradley, M. M. & Cuthbert, B. N. (1998). Emotion, motivation, and anxiety: Brain mechanisms and psychophysiology. *Biological psychiatry*, 44(12), 1248–1263 (siehe S. 70).
- Lara, O. D. & Labrador, M. A. (2012). A survey on human activity recognition using wearable sensors. *IEEE communications surveys & tutorials*, 15(3), 1192–1209 (siehe S. 178).
- Laref, R., Losson, E., Sava, A. & Siadat, M. (2018). Support Vector Machine Regression for Calibration Transfer between Electronic Noses Dedicated to Air Pollution Monitoring. *Sensors*, 18(11), 3716 (siehe S. 145, 157, 158).
- Larson, E. B., Wang, L., Bowen, J. D., McCormick, W. C., Teri, L., Crane, P. & Kukull, W. (2006). Exercise is associated with reduced risk for incident dementia among persons 65 years of age and older. *Annals of internal medicine*, 144(2), 73–81 (siehe S. 209).
- Latif, S., Rana, R., Younis, S., Qadir, J. & Epps, J. (2018). Transfer learning for improving speech emotion classification accuracy. *Interspeech 2018: Proceedings*, 257–261 (siehe S. 62).
- Lawanot, W., Inoue, M., Yokemura, T., Mongkolnam, P. & Nukoolkit, C. (2019). Daily stress and mood recognition system using deep learning and fuzzy clustering for promoting better well-being. *2019 IEEE International Conference on Consumer Electronics (ICCE)*, 1–6 (siehe S. 67).
- Leite, I., Pereira, A., Castellano, G., Mascarenhas, S., Martinho, C. & Paiva, A. (2011). Modelling empathy in social robotic companions. *International Conference on User Modeling, Adaptation, and Personalization*, 135–147 (siehe S. 182).
- Leonardi, C., Albertini, A., Pianesi, F. & Zancanaro, M. (2010). An exploratory study of a touch-based gestural interface for elderly. *Proceedings of the 6th nordic conference on human-computer interaction: Extending boundaries*, 845–850 (siehe S. 225).
- Lessel, P., Altmeyer, M., Kerber, F., Barz, M., Leidinger, C. & Krüger, A. (2016). WaterCoaster: A Device to Encourage People in a Playful Fashion to Reach Their Daily Water Intake Level. *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 1813–1820 (siehe S. 162, 165, 167).
- Leung, R., Tang, C., Haddad, S., McGrenere, J., Graf, P. & Ingriany, V. (2012). How older adults learn to use mobile devices: Survey and field investigations. *ACM Transactions on Accessible Computing (TACCESS)*, 4(3), 11 (siehe S. 205, 206).
- Lewis, D. D. & Gale, W. A. (1994). A sequential algorithm for training text classifiers. *SIGIR'94*, 3–12 (siehe S. 176).
- Liddell, C. & Guiney, C. (2015). Living in a cold and damp home: frameworks for understanding impacts on mental well-being. *Public Health*, 129(3), 191–199 (siehe S. 100).
- Lim, B. Y., Chng, X. & Zhao, S. (2017). Trade-off between automation and accuracy in mobile photo recognition food logging. *Proceedings of the Fifth International Symposium of Chinese CHI*, 53–59 (siehe S. 161, 174).
- Lin, Y.-P. & Jung, T.-P. (2017). Improving EEG-based emotion classification using conditional transfer learning. *Frontiers in human neuroscience*, 11, 334 (siehe S. 62).
- Lin, Y., Jessurun, J., de Vries, B. & Timmermans, H. (2011). Motivate: context aware mobile application for activity recommendation. *International Joint Conference on Ambient Intelligence*, 210–214 (siehe S. 205).
- Lindley, S. E., Harper, R. & Sellen, A. (2009). Desiring to be in touch in a changing communications landscape: attitudes of older adults. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1693–1702 (siehe S. 206).
- Lingenfelser, F., Wagner, J., André, E., McKeown, G. & Curran, W. (2014). An event driven fusion approach for enjoyment recognition in real-time. *Proceedings of the 22nd ACM international conference on Multimedia*, 377–386 (siehe S. 80, 81, 82).
- Liu, C., Conn, K., Sarkar, N. & Stone, W. (2008). Online affect detection and robot behavior adaptation for intervention of children with autism. *IEEE transactions on robotics*, 24(4), 883–896 (siehe S. 182, 243).
- Logan, B. (2000). Mel-Frequency Cepstral Coefficients for Music Modeling. *Proc. Int. Conf. on Music Information Retrieval (ISMIR)*, Plymouth, Massachusetts, 2000 (siehe S. 80).

- Loke, L. & Robertson, T. (2013). Moving and making strange: An embodied approach to movement-based interaction design. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 20(1), 1–25 (siehe S. 85).
- Louafi, S., Abdou, S. & Reiter, S. (2017). Effect of vegetation cover on thermal and visual comfort of pedestrians in urban spaces in hot and dry climate. *Nature & Technology*, (17), 30B (siehe S. 100).
- Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., Zhang, F., Chang, C.-L., Yong, M. G., Lee, J. et al. (2019). Mediapipe: A framework for building perception pipelines. *arXiv preprint arXiv:1906.08172* (siehe S. 40).
- Lumsden, J., Edwards, E. A., Lawrence, N. S., Coyle, D. & Munafò, M. R. (2016). Gamification of cognitive assessment and cognitive training: a systematic review of applications and efficacy. *JMIR serious games*, 4(2), e11 (siehe S. 236).
- Lutze, R., Baldauf, R. & Waldhör, K. (2015). Dehydration prevention and effective support for the elderly by the use of smartwatches. *2015 17th International Conference on E-health Networking, Application & Services (HealthCom)*, 404–409 (siehe S. 162).
- Ma, Z., Luo, G., Qin, K., Wang, N. & Niu, W. (2018). Online Sensor Drift Compensation for E-Nose Systems Using Domain Adaptation and Extreme Learning Machine. *Sensors*, 18(3). <https://doi.org/10.3390/s18030742> (siehe S. 157).
- Magnini, B., Balaraman, V., Dragoni, M., Guerini, M., Magnolini, S. & Piccioni, V. (2018). Ch1: A conversational system to calculate carbohydrates in a meal. *International Conference of the Italian Association for Artificial Intelligence*, 110–122 (siehe S. 190).
- Mankoff, J., Dey, A. K., Hsieh, G., Kientz, J., Lederer, S. & Ames, M. (2003). Heuristic evaluation of ambient displays. *Proceedings of the SIGCHI conference on Human factors in computing systems*, 169–176 (siehe S. 103, 135, 223).
- Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S. J. & McClosky, D. (2014). The Stanford CoreNLP Natural Language Processing Toolkit. *Association for Computational Linguistics (ACL) System Demonstrations*, 55–60 (siehe S. 192).
- Manor, B. R. & Gordon, E. (2003). Defining the temporal threshold for ocular fixation in free-viewing visuocognitive tasks. *Journal of neuroscience methods*, 128(1-2), 85–93 (siehe S. 75).
- Marin, J., Biswas, A., Ofli, F., Hynes, N., Salvador, A., Aytar, Y., Weber, I. & Torralba, A. (2019). Recipe1m+: A dataset for learning cross-modal embeddings for cooking recipes and food images. *IEEE transactions on pattern analysis and machine intelligence* (siehe S. 144).
- Marques, G., Ferreira, C. R. & Pitarma, R. (2019). Indoor Air Quality Assessment Using a CO₂ Monitoring System Based on Internet of Things. *Journal of medical systems*, 43(3), 67 (siehe S. 128).
- Martin, K., McLeod, E., Périard, J., Rattray, B., Keegan, R. & Pyne, D. B. (2019). The impact of environmental stress on cognitive performance: a systematic review. *Human factors*, 0018720819839817 (siehe S. 99).
- Martín, A., Lara-Cabrera, R., Fuentes-Hurtado, F., Naranjo, V. & Camacho, D. (2018). EvoDeep: A new evolutionary approach for automatic Deep Neural Networks parametrisation. *Journal of Parallel and Distributed Computing*, 117, 180–191 (siehe S. 106).
- Mateevitsi, V., Reda, K., Leigh, J. & Johnson, A. (2014). The health bar: a persuasive ambient display to improve the office worker's well being. *Proceedings of the 5th augmented human international conference* (siehe S. 121, 130).
- Matviienko, A., Rauschenberger, M., Cobus, V., Timmermann, J., Fortmann, J., Löcken, A., Müller, H., Trappe, C., Heuten, W. & Boll, S. (2015). Towards new ambient light systems: a close look at existing encodings of ambient light systems. *Interaction Design and Architecture (s)*. 2015;(26): 10-24. (siehe S. 121, 129).
- Matviienko, A., Rauschenberger, M., Cobus, V., Timmermann, J., Müller, H., Fortmann, J., Löcken, A., Trappe, C., Heuten, W. & Boll, S. (2015). Deriving design guidelines for ambient light systems. *Proceedings of the 14th international conference on mobile and ubiquitous multimedia*, 267–277 (siehe S. 121, 129).
- Maula, H., Hongisto, V., Naatula, V., Haapakangas, A. & Koskela, H. (2017). The effect of low ventilation rate with elevated bioeffluent concentration on work performance, perceived indoor air quality, and health symptoms. *Indoor air*, 27(6), 1141–1153 (siehe S. 97).
- Mauss, I. B., Shallcross, A. J., Troy, A. S., John, O. P., Ferrer, E., Wilhelm, F. H. & Gross, J. J. (2011). Don't hide your happiness! Positive emotion dissociation, social connectedness, and psychological functioning. *Journal of personality and social psychology*, 100(4), 738 (siehe S. 66).
- McFee, B., Raffel, C., Liang, D., Ellis, D. P., McVicar, M., Battenberg, E. & Nieto, O. (2015). librosa: Audio and music signal analysis in python. *Proceedings of the 14th python in science conference*, 8 (siehe S. 111).
- McGraw, I., Prabhavalkar, R., Alvarez, R., Arenas, M. G., Rao, K., Rybach, D., Alsharif, O., Sak, H., Gruenstein, A., Beaufays, F. et al. (2016). Personalized speech recognition on mobile devices. *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 5955–5959 (siehe S. 167).
- McGregor, G. R. & Vanos, J. K. (2018). Heat: a primer for public health researchers. *Public health*, 161, 138–146 (siehe S. 99).
- McLean, G. & Osei-Frimpong, K. (2019). Hey Alexa... examine the variables influencing the use of artificial intelligent in-home voice assistants. *Computers in Human Behavior*, 99, 28–37 (siehe S. 8, 197).
- Mendes, L., Ogink, N., Edouard, N., van Dooren, H., Tinôco, I. & Mosquera, J. (2015). NDIR gas sensor for spatial monitoring of carbon dioxide concentrations in naturally ventilated livestock buildings. *Sensors*, 15(5), 11239–11257 (siehe S. 31).
- Mesaros, A., Heittola, T. & Virtanen, T. (2018). A multi-device dataset for urban acoustic scene classification. *arXiv preprint arXiv:1807.09840* (siehe S. 110).

- Meschede, D. (2015). *Gerthsen Physik* (25. Auflage). Springer Spektrum. <https://doi.org/10.1007/978-3-662-45977-5>. (Siehe S. 32)
- Meyer, J., Arnrich, B., Schumm, J. & Troster, G. (2010). Design and modeling of a textile pressure sensor for sitting posture classification. *IEEE Sensors Journal*, 10(8), 1391–1398 (siehe S. 71).
- Michaelson, J., Abdallah, S., Steuer, N., Thompson, S., Marks, N., Akeh, J., Cordon, C. & Potts, R. (2009). National accounts of well-being: Bringing real wealth onto the balance sheet (siehe S. 13, 14).
- Milde, B. & Köhn, A. (2018). Open Source Automatic Speech Recognition for German. *Speech Communication; 13th ITG-Symposium*, 1–5 (siehe S. 192, 194).
- Miller, G. F. (1989). Designing Neural Networks using Genetic Algorithms. (siehe S. 106).
- Ming, Z.-Y., Chen, J., Cao, Y., Forde, C., Ngo, C.-W. & Chua, T. S. (2018). Food photo recognition for dietary tracking: System and experiment. *International Conference on Multimedia Modeling*, 129–141 (siehe S. 143).
- Mithbavkar, S. A. & Shah, M. S. (2019). Recognition of Emotion Through Facial Expressions Using EMG Signal. *2019 International Conference on Nascent Technologies in Engineering (ICNTE)*, 1–6 (siehe S. 67).
- Miu, T., Missier, P. & Plötz, T. (2015). Bootstrapping personalised human activity recognition models using online active learning. *2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing*, 1138–1147 (siehe S. 175, 176).
- Mokhayeri, F., Akbarzadeh-T, M. & Toosizadeh, S. (2011). Mental stress detection using physiological signals based on soft computing techniques. *2011 18th Iranian Conference of Biomedical Engineering (ICBME)*, 232–237 (siehe S. 75).
- Montesinos, V., Dell’Agnola, F., Arza, A., Aminifar, A. & Atienza, D. (2019). Multi-Modal Acute Stress Recognition Using Off-the-Shelf Wearable Devices. *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2196–2201 (siehe S. 67).
- Morris, M. E., Adair, B., Miller, K., Ozanne, E., Hansen, R., Pearce, A. J., Santamaria, N., Viegas, L., Long, M. & Said, C. M. (2013). Smart-home technologies to assist older people to live well at home. *Journal of aging science*, 1(1), 1–9 (siehe S. 207).
- Morse, G. & Stanley, K. O. (2016). Simple Evolutionary Optimization Can Rival Stochastic Gradient Descent in Neural Networks. *Proceedings of the Genetic and Evolutionary Computation Conference 2016*, 477–484. <https://doi.org/10.1145/2908812.2908916> (siehe S. 106).
- Motti, L. G., Vigouroux, N. & Gorce, P. (2013). Interaction techniques for older adults using touchscreen devices: a literature review. *Proceedings of the 25th Conference on l’Interaction Homme-Machine*, 125 (siehe S. 197, 206).
- Müller, H., Kazakova, A., Heuten, W. & Boll, S. (2015). Lighten Up! - An Ambient Light Progress Bar Using Individually Controllable LEDs. *European Conference on Ambient Intelligence*, 109–124 (siehe S. 130).
- Müller, H., Kazakova, A., Pielot, M., Heuten, W. & Boll, S. (2013). Ambient timer—unobtrusively reminding users of upcoming tasks with ambient light. *IFIP Conference on Human-Computer Interaction* (siehe S. 121, 130).
- Mureşan, H. & Oltean, M. (2018). Fruit recognition from images using deep learning. *Acta Universitatis Sapientiae, Informatica*, 10(1), 26–42 (siehe S. 144).
- Murnane, E. L., Huffaker, D. & Kossinets, G. (2015). Mobile health apps: adoption, adherence, and abandonment. *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers*, 261–264 (siehe S. 161, 163).
- Nakashima, Y., Kim, J., Flutura, S., Seiderer, A. & André, E. (2016). Stress Recognition in Daily Work. In S. Serino, A. Matic, D. Giakoumis, G. Lopez & P. Cipresso (Hrsg.), *Pervasive Computing Paradigms for Mental Health* (S. 23–33). Springer International Publishing. https://doi.org/10.1007/978-3-319-32270-4_3. (Siehe S. 70, 261, 264)
- Nakayoshi, M., Kanda, M., Shi, R. & de Dear, R. (2015). Outdoor thermal physiology along human pathways: a study using a wearable measurement system. *International journal of biometeorology*, 59(5), 503–515 (siehe S. 114).
- Nasr, M., Shokri, R. & Houmansadr, A. (2019). Comprehensive privacy analysis of deep learning: Passive and active white-box inference attacks against centralized and federated learning. *2019 IEEE Symposium on Security and Privacy (SP)*, 739–753 (siehe S. 6).
- Nazemi, H., Joseph, A., Park, J. & Emadi, A. (2019). Advanced Micro-and Nano-Gas Sensor Technology: A Review. *Sensors*, 19(6), 1285 (siehe S. 30, 31, 35).
- Neufert, E. & Neufert, P. (2012). *Architects’ data*. John Wiley & Sons. (Siehe S. 99).
- Nodine, C., Kundel, H., Toto, L. & Krupinski, E. A. (1992). Recording and analyzing eye-position data using a microcomputer workstation. *Behavior Research Methods, Instruments, & Computers*, 24(3), 475–485 (siehe S. 75).
- Núñez-Pacheco, C. & Loke, L. (2014). Aesthetic resources for technology-mediated bodily self-reflection: the case of eloquent robes. *Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: the Future of Design*, 1–10 (siehe S. 84).
- Nyklíček, I. & Kuijpers, K. F. (2008). Effects of mindfulness-based stress reduction intervention on psychological well-being and quality of life: is increased mindfulness indeed the mechanism? *Annals of Behavioral Medicine*, 35(3), 331–340 (siehe S. 83).
- Obermair, C., Reitberger, W., Meschtscherjakov, A., Lankes, M. & Tscheligi, M. (2008). perFrames: Persuasive picture frames for proper posture. *International conference on persuasive technology*, 128–139 (siehe S. 206).
- Occhialini, V., Van Essen, H. & Eggen, B. (2011). Design and evaluation of an ambient display to support time management during meetings. *IFIP Conference on Human-Computer Interaction*, 263–280 (siehe S. 121, 130).

- Onyeulo, E. B. & Gandhi, V. (2020). What makes a social robot good at interacting with humans? *Information*, 11(1), 43 (siehe S. 146).
- Ortiz Perez, A., Bierer, B., Scholz, L., Wöllenstein, J. & Palzer, S. (2018). A Wireless Gas Sensor Network to Monitor Indoor Environmental Quality in Schools. *Sensors*, 18(12), 4345 (siehe S. 129).
- Oviatt, S., Schuller, B., Cohen, P., Sonntag, D., Potamianos, G. & Krüger, A. (2019). *The Handbook of Multimodal-Multisensor Interfaces, Volume 3: Language Processing, Software, Commercialization, and Emerging Directions*. Morgan & Claypool. (Siehe S. 6).
- Ozdemir, M. A., Degirmenci, M., Guren, O. & Akan, A. (2019). EEG based Emotional State Estimation using 2-D Deep Learning Technique. *2019 Medical Technologies Congress (TIPTKNO)*, 1–4 (siehe S. 67).
- Ozdemir, M. A., Elagoz, B., Alaybeyoglu, A., Sadighzadeh, R. & Akan, A. (2019). Real Time Emotion Recognition from Facial Expressions Using CNN Architecture. *2019 Medical Technologies Congress (TIPTKNO)*, 1–4 (siehe S. 67).
- Pan, S. J. & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10), 1345–1359. <https://doi.org/10.1109/tkde.2009.191> (siehe S. 49, 50, 51, 52, 53)
- Parisi, G. I., Kemker, R., Part, J. L., Kanan, C. & Wermter, S. (2019). Continual lifelong learning with neural networks: A review. *Neural Networks*, 113, 54–71 (siehe S. 62).
- Park, B. J., Tsunetsugu, Y., Kasetani, T., Kagawa, T. & Miyazaki, Y. (2010). The physiological effects of Shinrin-yoku (taking in the forest atmosphere or forest bathing): evidence from field experiments in 24 forests across Japan. *Environmental health and preventive medicine*, 15(1), 18 (siehe S. 113).
- Peguero, E., Labrador, M. & Cook, B. (2016). Assessing jitter in sensor time series from android mobile devices. *2016 IEEE International Conference on Smart Computing (SMARTCOMP)*, 1–8 (siehe S. 40).
- Peterka, R. J. & Loughlin, P. J. (2004). Dynamic regulation of sensorimotor integration in human postural control. *Journal of neurophysiology*, 91(1), 410–423 (siehe S. 74).
- Picard, R. W. (2000). *Affective computing*. (Siehe S. 67).
- Piedrahita, R., Xiang, Y., Masson, N., Ortega, J., Collier, A., Jiang, Y., Li, K., Dick, R. P., Lv, Q., Hannigan, M. et al. (2014). The next generation of low-cost personal air quality sensors for quantitative exposure monitoring. *Atmospheric Measurement Techniques*, 7(10), 3325–3336 (siehe S. 32).
- Plarre, K., Raij, A., Hossain, S. M., Ali, A. A., Nakajima, M., Al'Absi, M., Ertin, E., Kamarck, T., Kumar, S., Scott, M. et al. (2011). Continuous inference of psychological stress from sensory measurements collected in the natural environment. *Proceedings of the 10th ACM/IEEE international conference on information processing in sensor networks*, 97–108 (siehe S. 71).
- Povey, D., Ghoshal, A., Boulianne, G., Burget, L., Glembek, O., Goel, N., Hannemann, M., Motlicek, P., Qian, Y., Schwarz, P. et al. (2011). The Kaldi speech recognition toolkit. *IEEE 2011 workshop on automatic speech recognition and understanding*, (CONF) (siehe S. 192).
- Prahm, C., Schulz, A., Paaßen, B., Schoisswohl, J., Kaniusas, E., Dorffner, G., Hammer, B. & Aszmann, O. (2019). Counteracting electrode shifts in upper-limb prosthesis control via transfer learning. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 27(5), 956–962 (siehe S. 62).
- Quan, T. (2016). *Molecular mechanisms of skin aging and age-related diseases*. CRC Press. (Siehe S. 204).
- Quinlan, J. R. (1993). *C4.5: Programs for Machine Learning*. Morgan Kaufmann. (Siehe S. 45).
- Rathmann, J., Beck, C., Flutura, S., Seiderer, A., Aslan, I. & André, E. (2020). Towards quantifying forest recreation: Exploring outdoor thermal physiology and human well-being along exemplary pathways in a central European urban forest (Augsburg, SE-Germany). *Urban Forestry & Urban Greening*, 49, 126622. <https://doi.org/https://doi.org/10.1016/j.ufug.2020.126622> (siehe S. 114, 261, 266)
- Rathmann, J. & Brumann, S. (2017). Therapeutische Landschaften in der Psychoonkologie: die gesundheitsfördernde Wirkung von Natur und Landschaft. *GAIA-Ecological Perspectives for Science and Society*, 26(3), 254–258 (siehe S. 113).
- Ravi, N., Dandekar, N., Mysore, P. & Littman, M. L. (2005). Activity recognition from accelerometer data. *Aaai*, 5(2005), 1541–1546 (siehe S. 178).
- Rayes, A. & Salam, S. (2019). *Internet of Things (IoT) Overview*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-99516-8>. (Siehe S. 5)
- Real, E., Moore, S., Selle, A., Saxena, S., Suematsu, Y. L., Tan, J., Le, Q. & Kurakin, A. (2017). Large-Scale Evolution of Image Classifiers (siehe S. 106).
- Ren, X., Yu, B., Lu, Y., Zhang, B., Hu, J. & Brombacher, A. (2019). LightSit: An unobtrusive health-promoting system for relaxation and fitness microbreaks at work. *Sensors*, 19(9), 2162 (siehe S. 79).
- Renaud, P. & Blondin, J.-P. (1997). The stress of Stroop performance: Physiological and emotional responses to color–word interference, task pacing, and pacing speed. *International Journal of Psychophysiology*, 27(2), 87–97 (siehe S. 71).
- Richter, G., Raban, D. R. & Rafaeli, S. (2015). Studying gamification: the effect of rewards and incentives on motivation. *Gamification in education and business* (S. 21–46). Springer. (Siehe S. 237).
- Ring, L., Shi, L., Totzke, K. & Bickmore, T. (2015). Social support agents for older adults: longitudinal affective computing in the home. *Journal on Multimodal User Interfaces*, 9(1), 79–88 (siehe S. 206).
- Rist, T., Seiderer, A. & André, E. (2018). Providing Life-Style-Intervention to Improve Well-Being of Elderly People. *International Conference on Entertainment Computing*, 362–367. https://doi.org/10.1007/978-3-319-99426-0_45 (siehe S. 202, 207, 226, 262, 264)

- Rist, T., **Seiderer, A.**, Hammer, S., Mayr, M. & André, E. (2015). CARE - extending a digital picture frame with a recommender mode to enhance well-being of elderly people. *2015 9th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth)*, 112–120. <https://doi.org/10.4108/icst.pervasivehealth.2015.259255> (siehe S. 13, 15, 202, 213, 221, 262, 263, 264)
- Ritschel, H., Aslan, I., Mertes, S., **Seiderer, A.** & André, E. (2019). Personalized synthesis of intentional and emotional non-verbal sounds for social robots. *8th International Conference on Affective Computing & Intelligent Interaction (ACII 2019)* (siehe S. 263, 265).
- Ritschel, H., Janowski, K., **Seiderer, A.** & André, E. (2019). Towards a Robotic Dietitian with Adaptive Linguistic Style. *Proceeding of the Poster and Workshop Sessions of Aml-2019, the 2019 European Conference on Ambient Intelligence* (siehe S. 263, 265).
- Ritschel, H., Janowski, K., **Seiderer, A.**, Wagner, S. & André, E. (2019). Insights on Usability and User Feedback for an Assistive Robotic Health Companion with Adaptive Linguistic Style. *Proceedings of the 12th ACM International Conference on Pervasive Technologies Related to Assistive Environments*, 319–320. <https://doi.org/10.1145/3316782.3322737> (siehe S. 202, 241, 246, 263, 265)
- Ritschel, H., **Seiderer, A.** & André, E. (2020). Pianobot: An Adaptive Robotic Piano Tutor. *Workshop on Exploring Creative Content in Social Robotics at HRI 2020*. (Siehe S. 263, 266).
- Ritschel, H., **Seiderer, A.**, Janowski, K., Aslan, I. & André, E. (2018). Drink-O-Mender: An Adaptive Robotic Drink Adviser. *Proceedings of the 3rd International Workshop on Multisensory Approaches to Human-Food Interaction*, 3. <https://doi.org/10.1145/3279954.3279957> (siehe S. 182, 183, 187, 262, 265)
- Ritschel, H., **Seiderer, A.**, Janowski, K., Wagner, S. & André, E. (2019). Adaptive Linguistic Style for an Assistive Robotic Health Companion Based on Explicit Human Feedback. *Proceedings of the 12th ACM International Conference on Pervasive Technologies Related to Assistive Environments*, 247–255. <https://doi.org/10.1145/3316782.3316791> (siehe S. 202, 241, 246, 263, 265)
- Rivas, I., Fussell, J. C., Kelly, F. J. & Querol, X. (2019). Indoor Sources of Air Pollutants. *Indoor Air Pollution* (S. 1–34). The Royal Society of Chemistry. <https://doi.org/10.1039/9781788016179-00001>. (Siehe S. 98)
- Roletscheck, C., Watzka, T., **Seiderer, A.**, Schiller, D. & André, E. Using an evolutionary approach to explore convolutional neural networks for acoustic scene classification. In: *Proceedings of the Detection and Classification of Acoustic Scenes and Events 2018 Workshop (DCASE2018)*. 2018, 158–162 (siehe S. 105, 261, 265).
- Roo, J. S., Gervais, R., Frey, J. & Hachet, M. (2017). Inner garden: Connecting inner states to a mixed reality sandbox for mindfulness. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 1459–1470 (siehe S. 83).
- Rowan, J. & Mynatt, E. D. (2005). Digital family portrait field trial: Support for aging in place. *Proceedings of the SIGCHI conference on Human factors in computing systems*, 521–530 (siehe S. 206).
- Rumelhart, D. E., Hinton, G. E. & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323, 533 (siehe S. 106).
- Runge, N., Samsonov, P., Degraen, D. & Schöning, J. (2016). No More Autobahn!: Scenic Route Generation Using Googles Street View. *Proceedings of the 21st International Conference on Intelligent User Interfaces*, 147–151 (siehe S. 113).
- Ryan, R. M. & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psychology*, 25(1), 54–67 (siehe S. 237).
- Sadek, I. & Biswas, J. (2019). Nonintrusive heart rate measurement using ballistocardiogram signals: a comparative study. *Signal, Image and Video Processing*, 13(3), 475–482 (siehe S. 78).
- Samsonov, P., Heller, F. & Schöning, J. (2017). Autobus: Selection of Passenger Seats Based on Viewing Experience for Touristic Tours. *Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia*, 321–326 (siehe S. 113).
- Sandler, M., Howard, A. G., Zhu, M., Zhmoginov, A. & Chen, L.-C. (2018). Inverted Residuals and Linear Bottlenecks: Mobile Networks for Classification, Detection and Segmentation. *CoRR, abs/1801.04381* (siehe S. 56).
- Satish, U., Mendell, M. J., Shekhar, K., Hotchi, T., Sullivan, D., Streufert, S. & Fisk, W. J. (2012). Is CO2 an indoor pollutant? Direct effects of low-to-moderate CO2 concentrations on human decision-making performance. *Environmental health perspectives*, 120(12), 1671 (siehe S. 97, 98).
- Sauvé, K., Houben, S., Marquardt, N., Bakker, S., Hengeveld, B., Gallacher, S. & Rogers, Y. (2017). LOOP: A physical artifact to facilitate seamless interaction with personal data in everyday life. *Proceedings of the 2017 ACM Conference Companion Publication on Designing Interactive Systems*, 285–288 (siehe S. 104).
- Savage, B. M., Lujan, H. L., Thipparathi, R. R. & DiCarlo, S. E. (2017). Humor, laughter, learning, and health! A brief review. *Advances in physiology education*, 41(3), 341–347 (siehe S. 68).
- Schafer, R. M. (1993). *The soundscape: Our sonic environment and the tuning of the world*. Simon; Schuster. (Siehe S. 101).
- Schiller, D., Huber, T., Lingenfelser, F., Dietz, M., **Seiderer, A.** & André, E. (2019). Relevance-based feature masking: Improving neural network based whale classification through explainable artificial intelligence. *20th Annual Conference of the International Speech Communication Association INTERSPEECH* (siehe S. 61, 261, 266).
- Schmidt, R. F. & Thews, G. (2013). *Physiologie des Menschen*. Springer-Verlag. (Siehe S. 102).
- Schneider, S., Goerlich, M. & Kummert, F. (2017). A framework for designing socially assistive robot interactions. *Cognitive Systems Research*, 43, 301–312 (siehe S. 243).

- Schroeder, V., Evans, E. D., Wu, Y.-C. M., Voll, C.-C. A., McDonald, B. R., Savagatrup, S. & Swager, T. M. (2019). Chemiresistive Sensor Array and Machine Learning Classification of Food. *ACS sensors*, 4(8), 2101–2108 (siehe S. 143).
- Schroeter, C., Mueller, S., Volkhardt, M., Einhorn, E., Huijnen, C., van den Heuvel, H., van Berlo, A., Bley, A. & Gross, H.-M. (2013). Realization and user evaluation of a companion robot for people with mild cognitive impairments. *2013 IEEE International Conference on Robotics and Automation*, 1153–1159 (siehe S. 242).
- Schuller, B., Steidl, S., Batliner, A., Hantke, S., Hönl, F., Orozco-Arroyave, J. R., Nöth, E., Zhang, Y. & Wengner, F. (2015). The INTERSPEECH 2015 computational paralinguistics challenge: nativeness, parkinson's & eating condition. *Sixteenth annual conference of the international speech communication association* (siehe S. 148).
- Schuller, D. M. & Schuller, B. W. (2020). The Challenge of Automatic Eating Behaviour Analysis and Tracking. In H. Costin, B. Schuller & A. M. Florea (Hrsg.), *Recent Advances in Intelligent Assistive Technologies: Paradigms and Applications* (S. 187–204). Springer International Publishing. https://doi.org/10.1007/978-3-030-30817-9_8. (Siehe S. 143)
- Schulze, F., Gao, X., Virzonis, D., Damiati, S., Schneider, M. & Kodzius, R. (2017). Air quality effects on human health and approaches for its assessment through microfluidic chips. *Genes*, 8(10), 244 (siehe S. 97, 99).
- Schwenke, T. (2016). *Private Nutzung von Smartglasses im öffentlichen Raum*. Fachinformationsdienst für internationale und interdisziplinäre ... (Siehe S. 143).
- Seiderer, A. & André, E. (2017). Development of a Multi-Device Nutrition Logging Prototype Including a Smartscale. *Proceedings of the 2017 International Conference on Digital Health - DH '17*, 239–240. <https://doi.org/10.1145/3079452.3079486> (siehe S. 160, 262, 264, 265)
- Seiderer, A., Aslan, I., Dang, C. T. & André, E. (2019). Indoor Air Quality and Wellbeing - Enabling Awareness and Sensitivity with Ambient IoT Displays. In I. Chatzigiannakis, B. De Ruyter & I. Mavrommati (Hrsg.), *Ambient Intelligence* (S. 266–282). Springer International Publishing. (Siehe S. 128, 262, 266).
- Seiderer, A., Dang, C. T. & André, E. (2017). Exploring Opportunistic Ambient Notifications in the Smart Home to Enhance Quality of Live. In M. Mokhtari, B. Abdulrazak & H. Aloulou (Hrsg.), *Enhanced Quality of Life and Smart Living* (S. 151–160). Springer International Publishing. https://doi.org/10.1007/978-3-319-66188-9_13. (Siehe S. 121, 262, 264)
- Seiderer, A., Dang, C. T. & André, E. (2019). Calibration Transfer for Food Recognition Models for E-Noses. *Proceedings of the 9th International Conference on the Internet of Things*, 36:1–36:4. <https://doi.org/10.1145/3365871.3365910> (siehe S. 157, 262, 266)
- Seiderer, A., Dietz, M., Aslan, I. & André, E. (2018). Enabling Privacy with Transfer Learning for Image Classification DNNs on Mobile Devices. *Proceedings of the 4th EAI International Conference on Smart Objects and Technologies for Social Good*, 25–30. <https://doi.org/10.1145/3284869.3284893> (siehe S. 54, 261, 265)
- Seiderer, A., Flutura, S. & André, E. (2017). Development of a mobile multi-device nutrition logger. *Proceedings of the 2nd ACM SIGCHI International Workshop on Multisensory Approaches to Human-Food Interaction*, 5–12. <https://doi.org/10.1145/3141788.3141790> (siehe S. 160, 262, 264, 265)
- Seiderer, A., Hammer, S., André, E., Mayr, M. & Rist, T. (2015). Exploring Digital Image Frames for Lifestyle Intervention to Improve Well-being of Older Adults. *Proceedings of the 5th International Conference on Digital Health 2015 - DH '15*, 71–78. <https://doi.org/10.1145/2750511.2750514> (siehe S. 202, 214, 221, 262, 263, 264)
- Seiderer, A., Ritschel, H. & André, E. (2020). Development of a Privacy-By-Design Speech Assistant Providing Nutrient Information for German Seniors. *Proceedings of the 6th EAI International Conference on Smart Objects and Technologies for Social Good*, 114–119. <https://doi.org/10.1145/3411170.3411227> (siehe S. 190, 262, 266)
- Settles, B. (2012). Active Learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 6(1), 1–114. <https://doi.org/10.2200/S00429ED1V01Y201207AIM018> (siehe S. 176)
- Seung, H. S., Oppor, M. & Sompolinsky, H. (1992). Query by committee. *Proceedings of the fifth annual workshop on Computational learning theory*, 287–294 (siehe S. 176).
- Shah, M. A., Shah, I. A., Lee, D.-G. & Hur, S. (2019). Design approaches of MEMS microphones for enhanced performance. *Journal of Sensors*, 2019 (siehe S. 27).
- Shahmohammadi, F., Hosseini, A., King, C. E. & Sarrafzadeh, M. (2017). Smartwatch based activity recognition using active learning. *2017 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*, 321–329 (siehe S. 146, 174, 175, 176).
- Shalini, S., Levins, T., Robinson, E. L., Lane, K., Park, G. & Skubic, M. (2019). Development and Comparison of Customized Voice-Assistant Systems for Independent Living Older Adults. *International Conference on Human-Computer Interaction*, 464–479 (siehe S. 147, 190, 196).
- Shapiro, S. L., Astin, J. A., Bishop, S. R. & Cordova, M. (2005). Mindfulness-based stress reduction for health care professionals: results from a randomized trial. *International journal of stress management*, 12(2), 164 (siehe S. 66, 83).
- Shinozawa, K., Naya, F., Yamato, J. & Kogure, K. (2005). Differences in effect of robot and screen agent recommendations on human decision-making. *International journal of human-computer studies*, 62(2), 267–279 (siehe S. 182).
- Shneiderman, B. (1982). The future of interactive systems and the emergence of direct manipulation. *Behaviour & Information Technology*, 1(3), 237–256. <https://doi.org/10.1080/01449298208914450> (siehe S. 146)
- Shoaib, M., Bosch, S., Incel, O. D., Scholten, H. & Havinga, P. J. (2016). Complex human activity recognition using smartphone and wrist-worn motion sensors. *Sensors*, 16(4), 426 (siehe S. 174, 176).

- Short, E., Swift-Spong, K., Greczek, J., Ramachandran, A., Litoiu, A., Grigore, E. C., Feil-Seifer, D., Shuster, S., Lee, J. J., Huang, S. et al. (2014). How to train your dragonbot: Socially assistive robots for teaching children about nutrition through play. *The 23rd IEEE international symposium on robot and human interactive communication*, 924–929 (siehe S. 182).
- Shu, L., Xie, J., Yang, M., Li, Z., Li, Z., Liao, D., Xu, X. & Yang, X. (2018). A review of emotion recognition using physiological signals. *Sensors*, 18(7), 2074 (siehe S. 67).
- Sidner, C. L., Bickmore, T., Nooraie, B., Rich, C., Ring, L., Shayganfar, M. & Vardoulakis, L. (2018). Creating new technologies for companionable agents to support isolated older adults. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 8(3), 17 (siehe S. 242).
- Skevington, S. M., Lotfy, M. & O'Connell, K. A. (2004). The World Health Organization's WHOQOL-BREF quality of life assessment: psychometric properties and results of the international field trial. A report from the WHOQOL group. *Quality of life Research*, 13(2), 299–310 (siehe S. 14).
- Slovák, P., Janssen, J. & Fitzpatrick, G. (2012). Understanding heart rate sharing: towards unpacking physiosocial space. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 859–868 (siehe S. 83, 84, 93).
- Song, C., Ikei, H., Igarashi, M., Miwa, M., Takagaki, M. & Miyazaki, Y. (2014). Physiological and psychological responses of young males during spring-time walks in urban parks. *Journal of physiological anthropology*, 33(1), 8 (siehe S. 114).
- Sonne, T. & Jensen, M. M. (2016). Chillfish: A respiration game for children with adhd. *Proceedings of the TEI'16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction*, 271–278 (siehe S. 83).
- Sørensen, H., Raptis, D., Kjeldskov, J. & Skov, M. B. (2014). The 4C framework: principles of interaction in digital ecosystems. *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 87–97 (siehe S. 143, 160, 164).
- Soubam, S., Agrawal, M. & Naik, V. (2017). Using an Arduino and a smartwatch to measure liquid consumed from any container. *2017 9th International Conference on Communication Systems and Networks (COMSNETS)*, 464–467 (siehe S. 162, 165, 167).
- Stanley, K. O. & Miikkulainen, R. (2002). Evolving neural networks through augmenting topologies. *Evolutionary computation*, 10(2), 99–127 (siehe S. 106).
- Steinke, F., Fritsch, T., Brem, D. & Simonsen, S. (2012). Requirement of AAL systems: older persons' trust in sensors and characteristics of AAL technologies. *Proceedings of the 5th International Conference on Pervasive Technologies Related to Assistive Environments*, 15 (siehe S. 204).
- Stevens, S. S., Volkman, J. & Newman, E. B. (1937). A scale for the measurement of the psychological magnitude pitch. *The Journal of the Acoustical Society of America*, 8(3), 185–190 (siehe S. 61).
- Stewart, I. D. & Oke, T. R. (2012). Local climate zones for urban temperature studies. *Bulletin of the American Meteorological Society*, 93(12), 1879–1900 (siehe S. 115).
- Stigall, B., Waycott, J., Baker, S. & Caine, K. (2019). Older Adults' Perception and Use of Voice User Interfaces: A Preliminary Review of the Computing Literature. *Proceedings of the 31st Australian Conference on Human-Computer-Interaction*, 423–427 (siehe S. 197).
- Sultanow, E. & Chircu, A. (2019). A Review of IoT Technologies, Standards, Tools, Frameworks and Platforms. *The Internet of Things in the Industrial Sector* (S. 3–34). Springer. (Siehe S. 17).
- Sutton, R. S. & Barto, A. G. (1998). Reinforcement Learning: An Introduction (siehe S. 45).
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. & Wojna, Z. (2015). Rethinking the Inception Architecture for Computer Vision. *CoRR, abs/1512.00567* (siehe S. 56).
- Tamura, T., Maeda, Y., Sekine, M. & Yoshida, M. (2014). Wearable photoplethysmographic sensors—past and present. *Electronics*, 3(2), 282–302 (siehe S. 21).
- Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C. & Liu, C. (2018). A survey on deep transfer learning. *International Conference on Artificial Neural Networks*, 270–279 (siehe S. 50, 53, 54, 62).
- Tao, V., Moy, K. & Amirfar, V. A. (2016). A little robot with big promise may be future of personalized health care. *Pharmacy Today*, 22(9), 38 (siehe S. 147, 182, 242).
- Tapus, A., Tapus, C. & Mataric, M. J. (2009). The use of socially assistive robots in the design of intelligent cognitive therapies for people with dementia. *2009 IEEE international conference on rehabilitation robotics*, 924–929 (siehe S. 243).
- Tapus, A., Țăpuș, C. & Matarić, M. J. (2008). User—robot personality matching and assistive robot behavior adaptation for post-stroke rehabilitation therapy. *Intelligent Service Robotics*, 1(2), 169 (siehe S. 182, 243).
- Tong, S. & Koller, D. (2001). Support vector machine active learning with applications to text classification. *Journal of machine learning research*, 2(Nov), 45–66 (siehe S. 176).
- Tsai, C.-H., Lin, C.-Y. & Lin, C.-J. (2014). Incremental and decremental training for linear classification. *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, 343–352 (siehe S. 178).
- Tsiakas, K., Abujelala, M. & Makedon, F. (2018). Task engagement as personalization feedback for socially-assistive robots and cognitive training. *Technologies*, 6(2), 49 (siehe S. 242).
- Tsiakas, K., Kyrarini, M., Karkaletsis, V., Makedon, F. & Korn, O. (2018). A Taxonomy in Robot-Assisted Training: Current Trends, Needs and Challenges. *Technologies*, 6(4), 119 (siehe S. 242).

- Tsujita, H. & Rekimoto, J. (2011). HappinessCounter: Smile-encouraging Appliance to Increase Positive Mood. *CHI '11 Extended Abstracts on Human Factors in Computing Systems*, 117–126. <https://doi.org/10.1145/1979742.1979608> (siehe S. 209)
- Turunen, M., Hakulinen, J., Smith, C., Charlton, D., Zhang, L. & Cavazza, M. (2008). Physically embodied conversational agents as health and fitness companions. *Ninth Annual Conference of the International Speech Communication Association* (siehe S. 206).
- Ulrich, R. S. (1984). View through a window may influence recovery from surgery. *Science*, 224(4647), 420–421 (siehe S. 113).
- Vapnik, V. (1998). *Statistical learning theory*. Wiley. (Siehe S. 52).
- Vardoulakis, L. P., Ring, L., Barry, B., Sidner, C. L. & Bickmore, T. (2012). Designing relational agents as long term social companions for older adults. *International Conference on Intelligent Virtual Agents*, 289–302 (siehe S. 205, 206).
- Vastenburg, M. H., Keyson, D. V. & de Ridder, H. (2009). Considerate Home Notification Systems: A User Study of Acceptability of Notifications in a Living-room Laboratory. *Int. J. Hum.-Comput. Stud.*, 67(9), 814–826. <https://doi.org/10.1016/j.ijhcs.2009.06.002> (siehe S. 121, 122)
- Verheijden Klompstra, L., Jaarsma, T. & Strömberg, A. (2014). Exergaming in older adults: a scoping review and implementation potential for patients with heart failure. *European Journal of Cardiovascular Nursing*, 13(5), 388–398 (siehe S. 236).
- Vi, C. T., Ablart, D., Arthur, D. & Obrist, M. (2017). Gustatory interface: the challenges of ‘how’ to stimulate the sense of taste. *Proceedings of the 2nd ACM SIGCHI International Workshop on Multisensory Approaches to Human-Food Interaction*, 29–33 (siehe S. 103).
- Vijayakumar, S., Flynn, R. & Murray, N. (2020). A comparative study of machine learning techniques for emotion recognition from peripheral physiological signals. *2020 31st Irish Signals and Systems Conference (ISSC)*, 1–6 (siehe S. 118).
- Vogt, T., André, E. & Bee, N. (2008). EmoVoice—A framework for online recognition of emotions from voice. *International Tutorial and Research Workshop on Perception and Interactive Technologies for Speech-Based Systems*, 188–199 (siehe S. 80).
- Wagner, J., Fraga-Silva, T., Josse, Y., Schiller, D., Seiderer, A. & André, E. (2017). Infected Phonemes: How a Cold Impairs Speech on a Phonetic Level. *INTERSPEECH*, 3457–3461 (siehe S. 263, 264).
- Wagner, J., Lingenfelser, F. & André, E. (2015). Building a robust system for multimodal emotion recognition. *Emotion recognition: A pattern analysis approach*, 379–410 (siehe S. 67).
- Wagner, J., Lingenfelser, F., Baur, T., Damian, I., Kistler, F. & André, E. (2013). The social signal interpretation (SSI) framework: multimodal signal processing and recognition in real-time. *Proceedings of the 21st ACM international conference on Multimedia*, 831–834 (siehe S. 40, 175).
- Wagner, J., Schiller, D., Seiderer, A. & André, E. (2018). Deep Learning in Paralinguistic Recognition Tasks: Are Hand-crafted Features Still Relevant? *Proc. Interspeech 2018*, 147–151 (siehe S. 263, 265).
- Wagner, J., Seiderer, A., Lingenfelser, F. & André, E. (2015). Combining Hierarchical Classification with Frequency Weighting for the Recognition of Eating Conditions. *INTERSPEECH*, 5 (siehe S. 148, 151, 262, 264).
- Walmink, W., Wilde, D. & Mueller, F. (2014). Displaying heart rate data on a bicycle helmet to support social exertion experiences. *Proceedings of the 8th International Conference on Tangible, Embedded and Embodied Interaction*, 97–104 (siehe S. 84).
- Ware, M., Frank, E., Holmes, G., Hall, M. & Witten, I. H. (2001). Interactive machine learning: letting users build classifiers. *International Journal of Human-Computer Studies*, 55(3), 281–292 (siehe S. 175).
- Weiser, M. (1991). The computer for the 21st century. *Scientific American*, 265(3), 94–104 (siehe S. 5, 103).
- Weiser, M. & Brown, J. S. (1995). Designing Calm Technology (siehe S. 103).
- Weiss, G. M., Timko, J. L., Gallagher, C. M., Yoneda, K. & Schreiber, A. J. (2016). Smartwatch-based activity recognition: A machine learning approach. *2016 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)*, 426–429 (siehe S. 143, 146).
- Weiss, K., Khoshgoftaar, T. M. & Wang, D. (2016). A survey of transfer learning. *Journal of Big data*, 3(1), 9 (siehe S. 4, 50, 62).
- Werner, J., Wettach, R. & Hornecker, E. (2008). United-pulse: feeling your partner’s pulse. *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services*, 535–538 (siehe S. 84).
- Werner, L., Huang, G. & Pitts, B. J. (2019). Automated Speech Recognition Systems and Older Adults: A Literature Review and Synthesis. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 63(1), 42–46 (siehe S. 197).
- Woodward, K., Kanjo, E., Burton, S. & Oikonomou, A. (2018). EmoEcho: A Tangible Interface to Convey and Communicate Emotions. *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*, 746–749 (siehe S. 84).
- Xie, Y., Liang, R., Liang, Z., Huang, C., Zou, C. & Schuller, B. (2019). Speech emotion classification using attention-based lstm. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 27(11), 1675–1685 (siehe S. 67).
- Yasavur, U., Lisetti, C. & Rishe, N. (2014). Let’s talk! speaking virtual counselor offers you a brief intervention. *Journal on Multimodal User Interfaces*, 8(4), 381–398 (siehe S. 206).
- Yim, J. (2016). Therapeutic benefits of laughter in mental health: a theoretical review. *The Tohoku journal of experimental medicine*, 239(3), 243–249 (siehe S. 66, 68, 209).

- Yu, B., Funk, M., Hu, J., Wang, Q. & Feijs, L. (2018). Biofeedback for everyday stress management: a systematic review. *Frontiers in ICT*, 5, 23 (siehe S. 69).
- Yu, B., Hu, J., Funk, M. & Feijs, L. (2017). A model of nature soundscape for calm information display. *Interacting with Computers*, 29(6), 813–823 (siehe S. 104).
- Yu, B., Hu, J., Funk, M. & Feijs, L. (2018). DeLight: biofeedback through ambient light for stress intervention and relaxation assistance. *Personal and Ubiquitous Computing*, 22(4), 787–805 (siehe S. 130).
- Yu, B., Zhang, B., An, P., Xu, L., Xue, M. & Hu, J. (2019). An Unobtrusive Stress Recognition System for the Smart Office. *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 1326–1329 (siehe S. 67).
- Zaharia, M., Xin, R. S., Wendell, P., Das, T., Armbrust, M., Dave, A., Meng, X., Rosen, J., Venkataraman, S., Franklin, M. J. et al. (2016). Apache spark: a unified engine for big data processing. *Communications of the ACM*, 59(11), 56–65 (siehe S. 42).
- Zakaria, R. A. (2010). NDIR instrumentation design for Carbon Dioxide gas sensing (siehe S. 32, 33).
- Zhai, J. & Barreto, A. (2006). Stress detection in computer users based on digital signal processing of noninvasive physiological variables. *2006 international conference of the IEEE engineering in medicine and biology society*, 1355–1358 (siehe S. 71, 78, 79).
- Zhai, J., Barreto, A. B., Chin, C. & Li, C. (2005). Realization of stress detection using psychophysiological signals for improvement of human-computer interactions. *Proceedings. IEEE SoutheastCon, 2005.*, 415–420 (siehe S. 74).
- Zhang, L., Tian, F.-C., Peng, X.-W. & Yin, X. (2014). A rapid discreteness correction scheme for reproducibility enhancement among a batch of MOS gas sensors. *Sensors and Actuators A: Physical*, 205, 170–176 (siehe S. 33).
- Zhang, X., Wargocki, P., Lian, Z. & Thyregod, C. (2017). Effects of exposure to carbon dioxide and bioeffluents on perceived air quality, self-assessed acute health symptoms, and cognitive performance. *Indoor air*, 27(1), 47–64 (siehe S. 97).
- Zhang, Z., Coutinho, E., Deng, J. & Schuller, B. (2014). Cooperative learning and its application to emotion recognition from speech. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 23(1), 115–126 (siehe S. 175).
- Zhou, B., Cheng, J., Sundholm, M., Reiss, A., Huang, W., Amft, O. & Lukowicz, P. (2015). Smart table surface: A novel approach to pervasive dining monitoring. *2015 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, 155–162 (siehe S. 143).
- Zimmerman, J., Forlizzi, J. & Evenson, S. (2007). Research through design as a method for interaction design research in HCI. *Proceedings of the SIGCHI conference on Human factors in computing systems*, 493–502 (siehe S. 85).
- Zuckerman, O. & Gal-Oz, A. (2014). Deconstructing gamification: evaluating the effectiveness of continuous measurement, virtual rewards, and social comparison for promoting physical activity. *Personal and ubiquitous computing*, 18(7), 1705–1719 (siehe S. 237).